

# CSC583-RonaldleeEjalu-HW#6

May 27, 2023

Ronaldlee Ejalu

CSC 583

HW#6

```
[1]: import sys
import torch
import numpy as np
```

## 1 Task 1

The task is to use the pipeline syntax to generate text:

### 1.0.1 1. Choose two ‘general’ models one small and one large.

```
[ ]: !pip install transformers
!pip install xformers
```

Looking in indexes: <https://pypi.org/simple>, <https://us-python.pkg.dev/colab-wheels/public/simple/>

Collecting transformers

Downloading transformers-4.29.2-py3-none-any.whl (7.1 MB)

↳ 7.1/7.1 MB

82.5 MB/s eta 0:00:00

Requirement already satisfied: filelock in /usr/local/lib/python3.10/dist-packages (from transformers) (3.12.0)

Collecting huggingface-hub<1.0,>=0.14.1 (from transformers)

Downloading huggingface\_hub-0.14.1-py3-none-any.whl (224 kB)

↳ 224.5/224.5 kB

27.7 MB/s eta 0:00:00

Requirement already satisfied: numpy>=1.17 in /usr/local/lib/python3.10/dist-packages (from transformers) (1.22.4)

Requirement already satisfied: packaging>=20.0 in  
/usr/local/lib/python3.10/dist-packages (from transformers) (23.1)  
Requirement already satisfied: pyyaml>=5.1 in /usr/local/lib/python3.10/dist-  
packages (from transformers) (6.0)  
Requirement already satisfied: regex!=2019.12.17 in  
/usr/local/lib/python3.10/dist-packages (from transformers) (2022.10.31)  
Requirement already satisfied: requests in /usr/local/lib/python3.10/dist-  
packages (from transformers) (2.27.1)  
Collecting tokenizers!=0.11.3,<0.14,>=0.11.1 (from transformers)  
 Downloading  
tokenizers-0.13.3-cp310-cp310-manylinux\_2\_17\_x86\_64.manylinux2014\_x86\_64.whl  
(7.8 MB)

↳  
↳7.8/7.8 MB  
112.5 MB/s eta 0:00:00

Requirement already satisfied: tqdm>=4.27 in  
/usr/local/lib/python3.10/dist-packages (from transformers) (4.65.0)  
Requirement already satisfied: fsspec in /usr/local/lib/python3.10/dist-packages  
(from huggingface-hub<1.0,>=0.14.1->transformers) (2023.4.0)  
Requirement already satisfied: typing-extensions>=3.7.4.3 in  
/usr/local/lib/python3.10/dist-packages (from huggingface-  
hub<1.0,>=0.14.1->transformers) (4.5.0)  
Requirement already satisfied: urllib3<1.27,>=1.21.1 in  
/usr/local/lib/python3.10/dist-packages (from requests->transformers) (1.26.15)  
Requirement already satisfied: certifi>=2017.4.17 in  
/usr/local/lib/python3.10/dist-packages (from requests->transformers)  
(2022.12.7)  
Requirement already satisfied: charset-normalizer~=2.0.0 in  
/usr/local/lib/python3.10/dist-packages (from requests->transformers) (2.0.12)  
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.10/dist-  
packages (from requests->transformers) (3.4)  
Installing collected packages: tokenizers, huggingface-hub, transformers  
Successfully installed huggingface-hub-0.14.1 tokenizers-0.13.3  
transformers-4.29.2

Looking in indexes: <https://pypi.org/simple>, [https://us-python.pkg.dev/colab-  
wheels/public/simple/](https://us-python.pkg.dev/colab-wheels/public/simple/)  
Collecting xformers

Downloading xformers-0.0.20-cp310-cp310-manylinux2014\_x86\_64.whl (109.1 MB)

↳  
↳109.1/109.1 MB  
16.8 MB/s eta 0:00:00

Requirement already satisfied: numpy in /usr/local/lib/python3.10/dist-  
packages (from xformers) (1.22.4)  
Collecting pyre-extensions==0.0.29 (from xformers)  
 Downloading pyre\_extensions-0.0.29-py3-none-any.whl (12 kB)

Requirement already satisfied: torch==2.0.1 in /usr/local/lib/python3.10/dist-packages (from xformers) (2.0.1+cu118)

Collecting typing-inspect (from pyre-extensions==0.0.29->xformers)

Downloading typing\_inspect-0.9.0-py3-none-any.whl (8.8 kB)

Requirement already satisfied: typing-extensions in /usr/local/lib/python3.10/dist-packages (from pyre-extensions==0.0.29->xformers) (4.5.0)

Requirement already satisfied: filelock in /usr/local/lib/python3.10/dist-packages (from torch==2.0.1->xformers) (3.12.0)

Requirement already satisfied: sympy in /usr/local/lib/python3.10/dist-packages (from torch==2.0.1->xformers) (1.11.1)

Requirement already satisfied: networkx in /usr/local/lib/python3.10/dist-packages (from torch==2.0.1->xformers) (3.1)

Requirement already satisfied: jinja2 in /usr/local/lib/python3.10/dist-packages (from torch==2.0.1->xformers) (3.1.2)

Requirement already satisfied: triton==2.0.0 in /usr/local/lib/python3.10/dist-packages (from torch==2.0.1->xformers) (2.0.0)

Requirement already satisfied: cmake in /usr/local/lib/python3.10/dist-packages (from triton==2.0.0->torch==2.0.1->xformers) (3.25.2)

Requirement already satisfied: lit in /usr/local/lib/python3.10/dist-packages (from triton==2.0.0->torch==2.0.1->xformers) (16.0.5)

Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.10/dist-packages (from jinja2->torch==2.0.1->xformers) (2.1.2)

Requirement already satisfied: mpmath>=0.19 in /usr/local/lib/python3.10/dist-packages (from sympy->torch==2.0.1->xformers) (1.3.0)

Collecting mypy-extensions>=0.3.0 (from typing-inspect->pyre-extensions==0.0.29->xformers)

Downloading mypy\_extensions-1.0.0-py3-none-any.whl (4.7 kB)

Installing collected packages: mypy-extensions, typing-inspect, pyre-extensions, xformers

Successfully installed mypy-extensions-1.0.0 pyre-extensions-0.0.29 typing-inspect-0.9.0 xformers-0.0.20

```
[ ]: from transformers import pipeline, set_seed
```

### openai-gpt

```
[ ]: generator = pipeline('text-generation', model='openai-gpt')
      set_seed(42)
      generator("Hello, HuggingFace,", max_length=30, num_return_sequences=3)
```

Some weights of OpenAIGPTLMHeadModel were not initialized from the model checkpoint at openai-gpt and are newly initialized: ['position\_ids']

You should probably TRAIN this model on a down-stream task to be able to use it for predictions and inference.

```
[ ]: [{'generated_text': 'Hello, HuggingFace, it\'s good to meet you. " \n " it\'s
good to meet you too. " huggingface'},
{'generated_text': "Hello, HuggingFace, what's this all about?\n'it's a small
affair ; a few guests are bringing up the food"},
{'generated_text': 'Hello, HuggingFace, " i said, trying to sound pleasant and
friendly. \n " if you have a place, i can give you'}]
```

### facebook/opt-1.3b

```
[ ]: generator = pipeline('text-generation', model="facebook/opt-1.3b",
↳do_sample=True, num_return_sequences=3)
set_seed(32)
gen = generator("Hello, I'm respectful and", max_length=30)
gen
```

```
[ ]: [{'generated_text': "Hello, I'm respectful and honest, willing to help out and
try and not be too bossy because I am an easy going person.\nCan"},
{'generated_text': "Hello, I'm respectful and I'll never take advantage. The
other two are toxic and I try and never talk to them again. If someone is"},
{'generated_text': "Hello, I'm respectful and interested in the trade. You're
interested in something off my list? Please share your list so I can have a
look"}]
```

## 1.0.2 2. Choose one 'domain specific' model: models pretrained on domain data

### bigscience/bloomz-560m

```
[ ]: generator = pipeline('text-generation', model="bigscience/bloomz-560m")
set_seed(32)
gen = generator("Translate to Spanish: good morning to everyone", max_length=30)
gen
```

```
Downloading (âĖ)lve/main/config.json: 0%|          | 0.00/715 [00:00<?, ?B/s]
Downloading pytorch_model.bin: 0%|          | 0.00/1.12G [00:00<?, ?B/s]
Downloading (âĖ)okenizer_config.json: 0%|          | 0.00/222 [00:00<?, ?B/s]
Downloading tokenizer.json: 0%|          | 0.00/14.5M [00:00<?, ?B/s]
Downloading (âĖ)cial_tokens_map.json: 0%|          | 0.00/85.0 [00:00<?, ?B/s]
```

```
[ ]: [{'generated_text': 'Translate to Spanish: good morning to everyone. Buenos dÃas
a todos.'}]
```

```
[ ]: # translating from French to English
gen = generator("Translate to English: comment allez-vous", max_length=30)
gen
```

```
[ ]: [{'generated_text': 'Translate to English: comment allez-vous? How are you?'}]
```

```
[ ]: [{'generated_text': 'Translate to English: Na, wie gehtâ€Žs it? Oh, how it is!'}]
```

```
[ ]: [{'generated_text': 'Translate to Telugu:como vai? àŕÕàřšàŕĲ  
↪àŕlāŕlārĲàŕŽàsĞārÿàsDārđāsAārCārèàŕf?'}]
```

```

        sys.exit(0)
    input_ids = tokenizer.encode(input_text, return_tensors='pt').to('cuda')
    attention_mask = torch.ones(input_ids.shape, dtype=torch.long, device='cuda')
    output = model.generate(input_ids, attention_mask=attention_mask,
        ↪max_length=512, do_sample=False, no_repeat_ngram_size=2, early_stopping=True)
    print(tokenizer.decode(output[0], skip_special_tokens=True))

```

```

Downloading (âĤ)olve/main/vocab.json: 0%|          | 0.00/1.04M [00:00<?, ?B/s]
Downloading (âĤ)olve/main/merges.txt: 0%|          | 0.00/456k [00:00<?, ?B/s]
Downloading (âĤ)olve/main/config.json: 0%|          | 0.00/666 [00:00<?, ?B/s]
Downloading pytorch_model.bin: 0%|          | 0.00/3.25G [00:00<?, ?B/s]
Downloading (âĤ)eneration_config.json: 0%|          | 0.00/124 [00:00<?, ?B/s]

```

```

<bound method ModuleUtilsMixin.num_parameters of GPT2LMHeadModel(
  (transformer): GPT2Model(
    (wte): Embedding(50257, 1280)
    (wpe): Embedding(1024, 1280)
    (drop): Dropout(p=0.1, inplace=False)
    (h): ModuleList(
      (0-35): 36 x GPT2Block(
        (ln_1): LayerNorm((1280,), eps=1e-05, elementwise_affine=True)
        (attn): GPT2Attention(
          (c_attn): Conv1D()
          (c_proj): Conv1D()
          (attn_dropout): Dropout(p=0.1, inplace=False)
          (resid_dropout): Dropout(p=0.1, inplace=False)
        )
        (ln_2): LayerNorm((1280,), eps=1e-05, elementwise_affine=True)
        (mlp): GPT2MLP(
          (c_fc): Conv1D()
          (c_proj): Conv1D()
          (act): NewGELUActivation()
          (dropout): Dropout(p=0.1, inplace=False)
        )
      )
    )
    (ln_f): LayerNorm((1280,), eps=1e-05, elementwise_affine=True)
  )
  (lm_head): Linear(in_features=1280, out_features=50257, bias=False)
)>

```

Type your input here:Bill Clinton

Setting `pad\_token\_id` to `eos\_token\_id`:50256 for open-end generation.

Bill Clinton's campaign manager, Patti Solis Doyle, said in a statement that the former president "has been a champion for women's rights and a tireless advocate for the rights of children."

"He has been an advocate and champion of women and children for decades," Doyle said. "He will be missed by all who knew him."

Type your input here:quit

An exception has occurred, use %tb to see the full traceback.

SystemExit: 0

/usr/local/lib/python3.10/dist-packages/IPython/core/interactiveshell.py:3561:

UserWarning: To exit: use 'exit', 'quit', or Ctrl-D.

warn("To exit: use 'exit', 'quit', or Ctrl-D.", stacklevel=1)

## 2.1 gpt2-large

### 2.1.1 Using the greedy search

```
[ ]: from transformers import GPT2Tokenizer, GPT2LMHeadModel

# load the tokenizer
tokenizer = GPT2Tokenizer.from_pretrained('gpt2-large')
model = GPT2LMHeadModel.from_pretrained('gpt2-large').to('cuda')
print(model.num_parameters)
sentinel = True
while sentinel:
    input_text = input("Type your input here: (Type exit or quit to quit the
→program)")
    if input_text == 'exit' or input_text == 'quit':
        sentinel = False
        exit(1)
    else:
        input_ids = tokenizer.encode(input_text, return_tensors='pt').to('cuda')
        attention_mask = torch.ones(input_ids.shape, dtype=torch.long,
→device='cuda')
        res = model.generate(input_ids, \
                             attention_mask=attention_mask, \
                             max_length=512
                             )
        for num, item in enumerate(res):
            print('%d : %s' %(num, tokenizer.decode(item, skip_special_tokens=True)))
```

Downloading (âĤ)olve/main/vocab.json: 0%| | 0.00/1.04M [00:00<?, ?B/s]

Downloading (âĤ)olve/main/merges.txt: 0%| | 0.00/456k [00:00<?, ?B/s]

Downloading (âĤ)olve/main/config.json: 0%| | 0.00/666 [00:00<?, ?B/s]

Downloading pytorch\_model.bin: 0%| | 0.00/3.25G [00:00<?, ?B/s]  
Downloading generation\_config.json: 0%| | 0.00/124 [00:00<?, ?B/s]

```
<bound method ModuleUtilsMixin.num_parameters of GPT2LMHeadModel(
  (transformer): GPT2Model(
    (wte): Embedding(50257, 1280)
    (wpe): Embedding(1024, 1280)
    (drop): Dropout(p=0.1, inplace=False)
    (h): ModuleList(
      (0-35): 36 x GPT2Block(
        (ln_1): LayerNorm((1280,), eps=1e-05, elementwise_affine=True)
        (attn): GPT2Attention(
          (c_attn): Conv1D()
          (c_proj): Conv1D()
          (attn_dropout): Dropout(p=0.1, inplace=False)
          (resid_dropout): Dropout(p=0.1, inplace=False)
        )
        (ln_2): LayerNorm((1280,), eps=1e-05, elementwise_affine=True)
        (mlp): GPT2MLP(
          (c_fc): Conv1D()
          (c_proj): Conv1D()
          (act): NewGELUActivation()
          (dropout): Dropout(p=0.1, inplace=False)
        )
      )
    )
    (ln_f): LayerNorm((1280,), eps=1e-05, elementwise_affine=True)
  )
  (lm_head): Linear(in_features=1280, out_features=50257, bias=False)
)>
```

Type your input here: (Type exit or quit to quit the program)Bill Clinton

Setting `pad\_token\_id` to `eos\_token\_id`:50256 for open-end generation.

0 : Bill Clinton's campaign manager, Patti Solis Doyle, said in a statement that the former president "has been a champion for women's rights and a champion for the rights of children."

"He has been a champion for the rights of children and families, and he will continue to be a champion for the rights of women and children," Doyle said.

Clinton's campaign also released a statement from his daughter, Chelsea Clinton, who said she was "honored" to be a part of the event.

"I am honored to be a part of the Clinton Foundation's Women's Leadership Initiative, which is dedicated to empowering women and girls around the world," Chelsea Clinton said. "I am also honored to be a part of the Clinton Global Initiative, which is dedicated to advancing the goals of the Clinton Foundation and advancing the interests of women and girls around the world."



The Clinton Foundation has been criticized for its lack of transparency, with critics saying it has not been transparent about its donors.

The foundation has said it has given away more than \$2 billion to women's charities since its inception in 2001.

The Clinton Foundation has also been criticized for its lack of transparency, with critics saying it has not been transparent about its donors.

The foundation has said it has given away more than \$2 billion to women's charities since its inception in 2001.

The Clinton Foundation has also been criticized for its lack of transparency, with critics saying it has not been transparent about its donors.

The foundation has said it has given away more than \$2 billion to women's charities since its inception in 2001.

The Clinton Foundation has also been criticized for its lack of transparency, with critics saying it has not been transparent about its donors.

The foundation has said it has given away more than \$2 billion to women's charities since its inception in 2001.

The Clinton Foundation has also been criticized for its lack of transparency, with critics saying it has not been transparent about its donors.

The foundation has said it has given away more than \$2 billion to women's charities since its inception in 2001.

The Clinton Foundation has also been criticized for its lack of transparency, with critics saying it has not been transparent about its donors.

The foundation has said it has given away more than \$2 billion to women's charities since its inception in 2001.

The Clinton Foundation has also been criticized for its lack of transparency, with critics saying it has not been

Type your input here: (Type exit or quit to quit the program)quit

One observation I am seeing with the greed search is repeatition of words

### 2.1.2 using the beam search

```
[ ]: from transformers import GPT2Tokenizer, GPT2LMHeadModel

# load the tokenizer
tokenizer = GPT2Tokenizer.from_pretrained('gpt2-large')
model = GPT2LMHeadModel.from_pretrained('gpt2-large').to('cuda')
print(model.num_parameters)
sentinel = True
while sentinel:
    input_text = input("Type your input here: (Type exit or quit to quit the
→program)")
    if input_text == 'exit' or input_text == 'quit':
        sentinel = False
        exit(1)
    else:
        input_ids = tokenizer.encode(input_text, return_tensors='pt').to('cuda')
        attention_mask = torch.ones(input_ids.shape, dtype=torch.long,
→device='cuda')
        res = model.generate(input_ids, \
                             attention_mask=attention_mask, \
                             max_length=512, \
                             num_beams = 5, \
                             num_return_sequences = 5, \
                             early_stopping=True
                             )
        for num, item in enumerate(res):
            print('%d : %s' %(num, tokenizer.decode(item, skip_special_tokens=True)))
```

```
<bound method ModuleUtilsMixin.num_parameters of GPT2LMHeadModel(
  (transformer): GPT2Model(
    (wte): Embedding(50257, 1280)
    (wpe): Embedding(1024, 1280)
    (drop): Dropout(p=0.1, inplace=False)
    (h): ModuleList(
      (0-35): 36 x GPT2Block(
        (ln_1): LayerNorm((1280,), eps=1e-05, elementwise_affine=True)
        (attn): GPT2Attention(
          (c_attn): Conv1D()
          (c_proj): Conv1D()
          (attn_dropout): Dropout(p=0.1, inplace=False)
          (resid_dropout): Dropout(p=0.1, inplace=False)
        )
        (ln_2): LayerNorm((1280,), eps=1e-05, elementwise_affine=True)
        (mlp): GPT2MLP(
          (c_fc): Conv1D()
          (c_proj): Conv1D()
```



Committee.

1 : Bill Clinton and his wife, former Secretary of State Hillary Rodham Clinton, leave the White House in Washington, D.C., U.S., on Tuesday, Oct. 11, 2016. (AP Photo/Carolyn Kaster)

WASHINGTON (AP) — The Latest on the 2016 presidential race (all times EDT):

9:30 p.m.

Hillary Clinton's campaign says it has raised more than \$1 million for the Democratic National Committee.

The campaign says in a statement that it has raised more than \$1 million for the Democratic National Committee.

The campaign says it has raised more than \$1 million for the Democratic National Committee.

The campaign says it has raised more than \$1 million for the Democratic National Committee.

The campaign says it has raised more than \$1 million for the Democratic National Committee.

The campaign says it has raised more than \$1 million for the Democratic National Committee.

The campaign says it has raised more than \$1 million for the Democratic National Committee.

The campaign says it has raised more than \$1 million for the Democratic National Committee.

The campaign says it has raised more than \$1 million for the Democratic National Committee.

2 : Bill Clinton and his wife, former Secretary of State Hillary Rodham Clinton, leave the White House in Washington, D.C., U.S., on Tuesday, Oct. 11, 2016. (AP Photo/Carolyn Kaster)

WASHINGTON (AP) — The Latest on the 2016 presidential race (all times EDT):

9:30 p.m.

Hillary Clinton's campaign says it has raised more than \$1 million for the Democratic National Committee.

The campaign says in a statement that it has raised more than \$1 million for the Democratic National Committee.

The campaign says it has raised more than \$1 million for the Democratic National Committee.

The campaign says it has raised more than \$1 million for the Democratic National Committee.

The campaign says it has raised more than \$1 million for the Democratic National Committee.

The campaign says it has raised more than \$1 million for the Democratic National Committee.

The campaign says it has raised more than \$1 million for the Democratic National Committee.

The campaign says it has raised more than \$1 million for the Democratic National Committee.

3 : Bill Clinton and his wife, former Secretary of State Hillary Rodham Clinton, leave the White House in Washington, D.C., U.S., on Tuesday, Oct. 11, 2016. (AP Photo/Carolyn Kaster)

WASHINGTON (AP) — The Latest on the 2016 presidential race (all times EDT):

9:30 p.m.

Hillary Clinton's campaign says it has raised more than \$1 million for the Democratic National Committee.

The campaign says in a statement that it has raised more than \$1 million for the Democratic National Committee.

The campaign says it has raised more than \$1 million for the Democratic National Committee.

The campaign says it has raised more than \$1 million for the Democratic National Committee.

The campaign says it has raised more than \$1 million for the Democratic National Committee.

The campaign says it has raised more than \$1 million for the Democratic National Committee.

The campaign says it has raised more than \$1 million for the Democratic National Committee.

4 : Bill Clinton and his wife, former Secretary of State Hillary Rodham Clinton, leave the White House in Washington, D.C., U.S., on Tuesday, Oct. 11, 2016. (AP

Photo/Carolyn Kaster)

WASHINGTON (AP) — The Latest on the 2016 presidential race (all times EDT):

9:30 p.m.

Hillary Clinton's campaign says it has raised more than \$1 million for the Democratic National Committee.

The campaign says in a statement that it has raised more than \$1 million for the Democratic National Committee.

The campaign says it has raised more than \$1 million for the Democratic National Committee.

The campaign says it has raised more than \$1 million for the Democratic National Committee.

The campaign says it has raised more than \$1 million for the Democratic National Committee.

The campaign says it has raised more than \$1 million for the Democratic National Committee.

Type your input here: (Type exit or quit to quit the program)exit

We still have the issue of repetition of the same words.

**we introduce n\_grams:** to prevent any repetition of the same words.

```
[ ]: from transformers import GPT2Tokenizer, GPT2LMHeadModel

# load the tokenizer
tokenizer = GPT2Tokenizer.from_pretrained('gpt2-large')
model = GPT2LMHeadModel.from_pretrained('gpt2-large').to('cuda')
print(model.num_parameters)
sentinel = True
while sentinel:
    input_text = input("Type your input here: (Type exit or quit to quit the
→program)")
    if input_text == 'exit' or input_text == 'quit':
        sentinel = False
        exit(1)
    else:
        input_ids = tokenizer.encode(input_text, return_tensors='pt').to('cuda')
        attention_mask = torch.ones(input_ids.shape, dtype=torch.long,
→device='cuda')
        res = model.generate(input_ids, \
```

```

        attention_mask=attention_mask, \
        max_length=512, \
        num_beams = 5, \
        num_return_sequences = 5, \
        no_repeat_ngram_size = 2, \
        early_stopping=True
    )
    for num, item in enumerate(res):
        print('%d : %s' %(num, tokenizer.decode(item, skip_special_tokens=True)))

```

```

<bound method ModuleUtilsMixin.num_parameters of GPT2LMHeadModel(
  (transformer): GPT2Model(
    (wte): Embedding(50257, 1280)
    (wpe): Embedding(1024, 1280)
    (drop): Dropout(p=0.1, inplace=False)
    (h): ModuleList(
      (0-35): 36 x GPT2Block(
        (ln_1): LayerNorm((1280,), eps=1e-05, elementwise_affine=True)
        (attn): GPT2Attention(
          (c_attn): Conv1D()
          (c_proj): Conv1D()
          (attn_dropout): Dropout(p=0.1, inplace=False)
          (resid_dropout): Dropout(p=0.1, inplace=False)
        )
        (ln_2): LayerNorm((1280,), eps=1e-05, elementwise_affine=True)
        (mlp): GPT2MLP(
          (c_fc): Conv1D()
          (c_proj): Conv1D()
          (act): NewGELUActivation()
          (dropout): Dropout(p=0.1, inplace=False)
        )
      )
    )
    (ln_f): LayerNorm((1280,), eps=1e-05, elementwise_affine=True)
  )
  (lm_head): Linear(in_features=1280, out_features=50257, bias=False)
)>

```

Type your input here: (Type exit or quit to quit the program)Bill Clinton

Setting `pad\_token\_id` to `eos\_token\_id`:50256 for open-end generation.

0 : Bill Clinton and his wife, former Secretary of State Hillary Rodham Clinton, leave the White House in Washington, D.C., U.S., on Tuesday, Oct. 11, 2016. (Photo: Andrew Harnik, AP)

WASHINGTON — President-elect Donald Trump's transition team said Tuesday that he will nominate retired Marine Gen. James "Mad Dog" Mattis to be secretary of defense, a position that would put him in charge of the nation's nuclear arsenal

and the military's fight against the Islamic State group in the Middle East.  
.@POTUS has chosen @MadDogMattis to lead the Department of Defense. â€” Donald J. Trump (@realDonaldTrump) December 2, 2015

, @realDonaldTrump has selected retired Army General James Mattis as his nominee to serve as our next SecDef. He is a man of tremendous talent and tremendous experience. <https://t.co/qZjVq6QjX1> â€” Kellyanne Conway (@KellyanneConway) November 29, 2017

The announcement comes as Trump prepares to take office on Jan. 20, and as the Republican-controlled Congress is set to begin work on legislation to repeal and replace the Affordable Care Act, also known as Obamacare, which Trump has vowed to replace with a plan he has called "insurance for everybody." The Senate is expected to vote on the legislation as early as next week, with the House likely to follow suit in early January.

1 : Bill Clinton and his wife, former Secretary of State Hillary Rodham Clinton, leave the White House in Washington, D.C., U.S., on Tuesday, Oct. 11, 2016.  
(Photo: Andrew Harnik, AP)

WASHINGTON â€” President-elect Donald Trump's transition team said Tuesday that he will nominate retired Marine Gen. James "Mad Dog" Mattis to be secretary of defense, a position that would put him in charge of the nation's nuclear arsenal and the military's fight against the Islamic State group in the Middle East.  
.@POTUS has chosen @MadDogMattis to lead the Department of Defense. â€” Donald J. Trump (@realDonaldTrump) December 2, 2015

, @realDonaldTrump has selected retired Army General James Mattis as his nominee to serve as our next SecDef. He is a man of tremendous talent and tremendous experience. <https://t.co/qZjVq6QjX1> â€” Kellyanne Conway (@KellyanneConway) November 29, 2017

The announcement comes as Trump prepares to take office on Jan. 20, and as the Republican-controlled Congress is set to begin work on legislation to repeal and replace the Affordable Care Act, also known as Obamacare, which Trump has vowed to replace with a plan he has called "insurance for everybody." The Senate is expected to vote on the legislation as early as next week, with the House likely to follow suit in early January, according to the New York Times.

2 : Bill Clinton and his wife, former Secretary of State Hillary Rodham Clinton, leave the White House in Washington, D.C., U.S., on Tuesday, Oct. 11, 2016.  
(Photo: Andrew Harnik, AP)

WASHINGTON â€” President-elect Donald Trump's transition team said Tuesday that he will nominate retired Marine Gen. James "Mad Dog" Mattis to be secretary of defense, a position that would put him in charge of the nation's nuclear arsenal and the military's fight against the Islamic State group in the Middle East.  
.@POTUS has chosen @MadDogMattis to lead the Department of Defense. â€” Donald J. Trump (@realDonaldTrump) December 2, 2015

, @realDonaldTrump has selected retired Army General James Mattis as his nominee to serve as our next SecDef. He is a man of tremendous talent and tremendous experience. <https://t.co/qZjVq6QjX1> â€” Kellyanne Conway (@KellyanneConway) November 29, 2017

The announcement comes as Trump prepares to take office on Jan. 20, and as the



Republican-controlled Congress is set to begin work on legislation to repeal and replace the Affordable Care Act, also known as Obamacare, which Trump has vowed to replace with a plan he has called "insurance for everybody." The Senate is expected to vote on the legislation as early as next week, with the House likely to follow suit in early January, according to the New York Times.

Read or Share this story: <http://usat.ly/2k0q9ZR>

3 : Bill Clinton and his wife, former Secretary of State Hillary Rodham Clinton, leave the White House in Washington, D.C., U.S., on Tuesday, Oct. 11, 2016.

(Photo: Andrew Harnik, AP)

WASHINGTON -- President-elect Donald Trump's transition team said Tuesday that he will nominate retired Marine Gen. James "Mad Dog" Mattis to be secretary of defense, a position that would put him in charge of the nation's nuclear arsenal and the military's fight against the Islamic State group in the Middle East.

.@POTUS has chosen @MadDogMattis to lead the Department of Defense. -- Donald J. Trump (@realDonaldTrump) December 2, 2015

, @realDonaldTrump has selected retired Army General James Mattis as his nominee to serve as our next SecDef. He is a man of tremendous talent and tremendous experience. <https://t.co/qZjVq6QjX1> -- Kellyanne Conway (@KellyanneConway) November 29, 2017

The announcement comes as Trump prepares to take office on Jan. 20, and as the Republican-controlled Congress is set to begin work on legislation to repeal and replace the Affordable Care Act, also known as Obamacare, which Trump has vowed to replace with a plan he has called "insurance for everybody." The Senate is expected to vote on the legislation as early as next week, with the House likely to follow suit in early January, according to the New York Times.

Read or Share this story: <http://usat.ly/2k0q9Qr>

4 : Bill Clinton and his wife, former Secretary of State Hillary Rodham Clinton, leave the White House in Washington, D.C., U.S., on Tuesday, Oct. 11, 2016.

(Photo: Andrew Harnik, AP)

WASHINGTON -- President-elect Donald Trump's transition team said Tuesday that he will nominate retired Marine Gen. James "Mad Dog" Mattis to be secretary of defense, a position that would put him in charge of the nation's nuclear arsenal and the military's fight against the Islamic State group in the Middle East.

.@POTUS has chosen @MadDogMattis to lead the Department of Defense. -- Donald J. Trump (@realDonaldTrump) December 2, 2015

, @realDonaldTrump has selected retired Army General James Mattis as his nominee to serve as our next SecDef. He is a man of tremendous talent and tremendous experience. <https://t.co/qZjVq6QjX1> -- Kellyanne Conway (@KellyanneConway) November 29, 2017

The announcement comes as Trump prepares to take office on Jan. 20, and as the Republican-controlled Congress is set to begin work on legislation to repeal and replace the Affordable Care Act, also known as Obamacare, which Trump has vowed

to replace with a plan he has called "insurance for everybody." The Senate is expected to vote on the legislation as early as next week, with the House likely to follow suit in early January, according to the New York Times.

Read or Share this story: <http://usat.ly/2k0q9Zr>

Type your input here: (Type exit or quit to quit the program)exit

repetitive sequences of sentences are generated

### 2.1.3 using top-k sampling

```
[ ]: from transformers import GPT2Tokenizer, GPT2LMHeadModel

# load the tokenizer
tokenizer = GPT2Tokenizer.from_pretrained('gpt2-large')
model = GPT2LMHeadModel.from_pretrained('gpt2-large').to('cuda')
print(model.num_parameters)
sentinel = True
while sentinel:
    input_text = input("Type your input here: (Type exit or quit to quit the_
→program)")
    if input_text == 'exit' or input_text == 'quit':
        sentinel = False
        exit(1)
    else:
        input_ids = tokenizer.encode(input_text, return_tensors='pt').to('cuda')
        attention_mask = torch.ones(input_ids.shape, dtype=torch.long,
→device='cuda')
        res = model.generate(input_ids, \
                             attention_mask=attention_mask, \
                             do_sample = True, \
                             max_length = 512, \
                             top_k = 0, \
                             #num_beams = 5, \
                             num_return_sequences = 5, \
                             #no_repeat_ngram_size = 2, \
                             #early_stopping = True
                             )
        for num, item in enumerate(res):
            print('%d : %s' %(num, tokenizer.decode(item, skip_special_tokens=True)))
```

```
<bound method ModuleUtilsMixin.num_parameters of GPT2LMHeadModel(
(transformer): GPT2Model(
  (wte): Embedding(50257, 1280)
  (wpe): Embedding(1024, 1280)
  (drop): Dropout(p=0.1, inplace=False)
```

```

(h): ModuleList(
  (0-35): 36 x GPT2Block(
    (ln_1): LayerNorm((1280,), eps=1e-05, elementwise_affine=True)
    (attn): GPT2Attention(
      (c_attn): Conv1D()
      (c_proj): Conv1D()
      (attn_dropout): Dropout(p=0.1, inplace=False)
      (resid_dropout): Dropout(p=0.1, inplace=False)
    )
    (ln_2): LayerNorm((1280,), eps=1e-05, elementwise_affine=True)
    (mlp): GPT2MLP(
      (c_fc): Conv1D()
      (c_proj): Conv1D()
      (act): NewGELUActivation()
      (dropout): Dropout(p=0.1, inplace=False)
    )
  )
)
(ln_f): LayerNorm((1280,), eps=1e-05, elementwise_affine=True)
)
(lm_head): Linear(in_features=1280, out_features=50257, bias=False)
)>
Type your input here: (Type exit or quit to quit the program)Bill Clinton
Setting `pad_token_id` to `eos_token_id`:50256 for open-end generation.

0 : Bill Clinton's email scandal is similar to the one Robert Reich wrote about
for the magazine in 1996 that led to his resignation. Clinton was under subpoena
  before and after the public release of what were then classified emails
between Clinton and Carlos Slim  to answer questions about whether he sent
classified information through a private personal account. There may or may not
be more Hillary Clinton emails, too.

At that point, and after the White House granted a date to a House Benghazi
committee lawyer to make public a portion of the Clinton archive on the grounds
that it would damage national security, the vice president became the consummate
public servant, sticking firmly to his story that he and he alone had answers
for Republicans in casting a curious line between appropriate executive power
and unlawful arming of terrorists.

And, as the nation mourns, in part because his opponent, Donald Trump, has
turned that strategy around by offering a different version of the narrative
over what he says is his own short time in public life, he may continue to pull
that precise shift.

"Next stop: President Trump," Frankfurt writes. He writes, "I'll be listening."
1 : Bill Clinton, though, will take his case to any stop line.

On his last day as president on Nov. 8, Bill Clinton will appeal a ruling in his

```

lawsuit against the university. "Those of you who bring these lawsuits should know from hearing the general counsel say, 'It wouldn't be worth your time pursuing this claim now,' " the plaintiff's lawyer, Philip Mitchell, told reporters last week in a brief court hearing. Mitchell has appealed to the 7th U.S. Circuit Court of Appeals to overturn the University of Arkansas's decision in August of 1998 to cut ties with its Clinton alumni. A hearing is set for late next month.

The case was the subject of an on-air humor skit staged by WKRF and WHAS 11. "The University of Arkansas is developing some hope in Jim Wirtz's spirit." WKRF called it "the greatest third shift ever!" Clinton Business Group Communications handled the leaving. A Catholic university, the university is not required to list a religious affiliation, so it lists a "religious mind-set," according to a press release in September 1998, which read, "It is the University's belief that everyone is made in the image of God."

As president, Bill Clinton vetoed a bill signed by Arkansas Governor Bill Clinton in 1993 to create the Center for Understanding Islamic Cultural Issues. The bill included musings on the lines of "Islamic law," as opposed to "law," with "Islam" being the sect. "The purpose of the center was to delve into issues in Islamic Law. However, when the bill became law, it became a problem for the University of Arkansas," the university president wrote in the journal Time. "It will not take part in the center's program or activities. The University determines what issues to cover and should not interfere with the religious policies of the country."

HECLA

33 Audacious University Invests in Sugar Marching in unusual manner, Jeremiah Donthea, CQ Independent Woman 06-11-01

"SMILE, Y'ALL," said William Tobias. This jovial Young and new guy and all around drum major in an Air Force Dance Company was pulling in the bucks at the money-accumulating WSU. "There's nothing better than to step into a drum set and become a participant in some alchemical ritual?" Tobias said, as the two-man Wednesday Express filled dance hall Vashon with the normality of summer's starts and ends. Outside

2 : Bill Clinton left office in 1993, a legacy that's captured the imagination of the nation. He was the champion of a generation of grieving widows -- philosophers and poets, parents of children killed regularly in action, they still fiercely pointed to him and to his Democratic leaders as heroes and showered them with ocean liners, flowers and handshakes.

They even founded a search-and-rescue organization: the National Whitewater Institute. They always, no matter what he was later implicated in -- and how he often makes a comeback anyway -- said lovingly of that presidency, that it represented a new place in America.

Hillary Clinton, meanwhile, made herself an undercard on all that late-deciding 2000s "moment" celebrity. She gained fame at a time when a good ol' girl from Indiana or Arkansas rang in an most noteworthy electoral anniversary on no less a national stage than in a city of somber necessity.     David Fahrenthold

Bill Clinton's approval ratings

View Graphic Fact check on Bill Clinton approval ratings

The workhorse of the Clinton administration of the late 1990s     whose foreign policy proved consequential for a generation     had a niche that was arguably narrower than many of his smaller-title associates. He had several strengths. Not the most eloquent but widespread, he was reassuringly predictable at a time when custodial responsibility for a nation he had just entered so glibly was not something a president could count on. He was persuasive and loved to tell it. And he had even more subtlety and style that helped calm zealots who would lash out at him. Those factors would come in handy during his six years as president. His job was to remind the nation of who he was     of his humanitarian side, his equation-making prowess and his halting capabilities of taking bipartisanship to the pitch of compromise. But those strengths came naturally to Clinton, too, just as they did to everyone else on the national stage. And with some of those strengths came responsibilities.

But he never got credit     not even though he left office 39 days after leaving office (one time was by shipwreck; one time a hard-luck existence of mental illness in a Miami electric chair) and with an approval rating of 55 percent to just 40 percent among GenXers 10 years out of college. Nor did the nation forget the same awkward incident where Bill left the office he shared with Hillary at her Sepuncio ranch in 1974 and retreated to the White House to have

3 : Bill Clinton was also featured in adocumentary film made by John Barry, a progressive filmmaker who opposed Bill Clinton in an earlier election. Barry reported that people in the area "were extremely depressed" during the campaign because of "articulations" of attacks on liberals from the Clinton campaign. Also, Barry quotes Dr. Helen Silva, a psychologist who gives psychological counseling to coffee shop employees and has advised Dr. Weiner to get help for this emotional night that has followed a Clinton victory. Silva had a heart conversation with Dr. Weiner after a Clinton win.

One man told Silva that Clinton obscured his interior voice. He told someone else, "He does not let no one talk."

CNN.com reported that several of Donald Trump's long-time friends and business associates had urged him to support Hillary Clinton following Election Day. Trump took to Twitter following the tweets and by Tuesday morning he called it a "sad day for our country."

4 : Bill Clinton is no Dick Cheney. For the first time, Republicans are not ignoring Clinton's record, where he praises a hero for invading Iraq, the surveillance state, way-big-un-then trade deals, and should I say, Obamacare.

The voters actually wanted someone who's first name was "Clinton," and he promised to be that. But now we get to see just how incompetent he really is, in one political speech after another, talked about his curling iron and how much he likes hockey. We get to see how Bill Clinton sent Katrina to Arkansas and says it was "a good thing." First, he said Katrina was a "bad thing." Then, somehow it was happening twice and it also generated 80,000 new jobs. Here is the great joke, which he recently spatted in contribution to that Wall Street Journal. In my view, Bill Clinton was President because he was able to spin a situation. A president can't meet a situation out of the blue, or down the road. As a lawyer, I don't talk about milliseconds from tomorrow, but Bill Clinton traveled to Sandy Hook on Sunday and at the funeral. He learned about the disaster, took the tragedy, and made it a 37-minute event. That makes him a decider on events, and that is how events and governors and mayors are won. Barack Obama and John Kerry had lots of speeches. But they only came to focus on life, not speech-making. Clinton compared banishing Sandy Hook to Lyndon Johnson dealing with the Vietnam War draft, a reference to a 15-year-old speech Clinton gave at Yale days after Kennedy's assassination, released in the waning days of the Kennedy presidency. Instead of it being a good thing, in a contextual sense, he pointed out that it wasn't sensible to have an undifferentiated draft at Halloween. But witness the many lawyers who must try to defend the Clinton speech-making. It was a mistake, somebody raised it. But now what does that tell us about who's listening? Clinton's over 25 years older than Obama is. They have a decade of experience in governing. If he could not make a run for the Oval Office, both will have to explain their mistakes over and over, and I wasn't impressed. Barack Obama has improved from his first election, but some of his fundamental flaws are still with us, purely for his own good. There is much we can learn from both - in both categories of critiquing. And maybe I would have voted for Al Gore had he become

Type your input here: (Type exit or quit to quit the program)quit

This looks much better it lacks consistency.

#### 2.1.4 using top-p (nucleus) sampling

```
[ ]: from transformers import GPT2Tokenizer, GPT2LMHeadModel

# load the tokenizer
tokenizer = GPT2Tokenizer.from_pretrained('gpt2-large')
model = GPT2LMHeadModel.from_pretrained('gpt2-large').to('cuda')
print(model.num_parameters)
sentinel = True
while sentinel:
    input_text = input("Type your input here: (Type exit or quit to quit the_
→program)")
    if input_text == 'exit' or input_text == 'quit':
        sentinel = False
    exit(1)
```

```

else:
    input_ids = tokenizer.encode(input_text, return_tensors='pt').to('cuda')
    attention_mask = torch.ones(input_ids.shape, dtype=torch.long,
    ↪device='cuda')
    res = model.generate(input_ids, \
                        attention_mask=attention_mask, \
                        do_sample = True, \
                        max_length = 512, \
                        top_k = 0, \
                        top_p = 0.7, \
                        #num_beams = 5, \
                        num_return_sequences = 3
                        )
    for num, item in enumerate(res):
        print('%d : %s' %(num, tokenizer.decode(item, skip_special_tokens=True)))

```

```

<bound method ModuleUtilsMixin.num_parameters of GPT2LMHeadModel(
  (transformer): GPT2Model(
    (wte): Embedding(50257, 1280)
    (wpe): Embedding(1024, 1280)
    (drop): Dropout(p=0.1, inplace=False)
    (h): ModuleList(
      (0-35): 36 x GPT2Block(
        (ln_1): LayerNorm((1280,), eps=1e-05, elementwise_affine=True)
        (attn): GPT2Attention(
          (c_attn): Conv1D()
          (c_proj): Conv1D()
          (attn_dropout): Dropout(p=0.1, inplace=False)
          (resid_dropout): Dropout(p=0.1, inplace=False)
        )
        (ln_2): LayerNorm((1280,), eps=1e-05, elementwise_affine=True)
        (mlp): GPT2MLP(
          (c_fc): Conv1D()
          (c_proj): Conv1D()
          (act): NewGELUActivation()
          (dropout): Dropout(p=0.1, inplace=False)
        )
      )
    )
    (ln_f): LayerNorm((1280,), eps=1e-05, elementwise_affine=True)
  )
  (lm_head): Linear(in_features=1280, out_features=50257, bias=False)
)>

```

Type your input here: (Type exit or quit to quit the program)Bill Clinton

Setting `pad\_token\_id` to `eos\_token\_id`:50256 for open-end generation.

0 : Bill Clinton was seen as a figure of change for women in America and the

Democratic Party, and in 1997 he began his second term as president, winning a second term in the White House.

1 : Bill Clinton would have appointed Jack G. Kennedy to be his Secretary of State and had the right to appoint his vice president.

You can't beat the simplicity of that.

You can also understand why he chose the left-wing socialist, G. William Domhoff, as his Secretary of State.

As you may know, Domhoff was one of the first major proponents of the welfare state, arguing that the state should help poor families make ends meet.

Domhoff is the same guy who has admitted he wants to abolish the minimum wage.

If he gets his way, the entire U.S. government will be in the hands of an unelected bureaucracy.

Dominghoff also pushed for an income tax and is a major advocate of gun control.

It was that, coupled with his socialist views, that got Domhoff the job of Secretary of State.

At least, that's the explanation as to why Bill Clinton nominated him.

But as you can see, Domhoff wasn't exactly a progressive.

What makes Domhoff even more disgusting is that he was the Chairman of the International Monetary Fund.

So you can imagine how this man felt when he was named Secretary of State.

He immediately flew to Japan to meet with the leaders of that country and agreed to take the lead on the global economy.

If that doesn't make you sick, I don't know what will.

This was before he had a chance to do his job and find out what kind of disaster was coming.

As you can see, he was very cooperative and was willing to meet with any country in the world.

He was also able to get Japan to implement the North American Free Trade Agreement.

But when the TPP went down in flames, Domhoff was completely unapologetic.



In fact, he was quite the pro-TPP patsy.

In fact, he even helped out with the whole process of getting it passed.

As you can see, he was quite willing to go along with anything to get this country off of the crapper.

He was a very good friend of the left-wing socialist, Margaret Thatcher.

In fact, she actually invited him to visit her country and try to fix it.

If he hadn't done that, then we would never have gotten the TPP passed.

But of course, he did and we never would have gotten  
2 : Bill Clinton's impeachment is another example of how the elites have been stealing from the middle class for the last 60 years.

You see, it's not just Bill Clinton who's been on the wrong side of the legal system. It's Hillary Clinton. As the head of the State Department, she allowed the crime of the century to occur in Haiti, and we all know how that turned out.

The State Department also played a role in the attack on Libya. Hillary Clinton also sided with the Saudis and the Qataris who were trying to overthrow Gaddafi, even though the U.S. was supplying arms to the rebels.

Then, of course, there's Benghazi. The State Department allowed the murder of four Americans, including Ambassador Christopher Stevens, to happen.

Finally, of course, there's Hillary Clinton's email scandal. It's an embarrassing affair that she has done nothing to hide.

So when the elites try to make it seem as though they are all on the same side, it's always a little bit of a charade.

And that's why we have to get rid of them.

Bernie Sanders is on the wrong side of this fight. He's not talking about destroying the big banks, the big oil companies or the wealthy elite, but he's talking about transforming America into a nation of the people, by the people and for the people.

Sanders wants to invest in our country and rebuild our country. He wants to help working families, and the working poor, and people of color, and seniors, and veterans. He wants to help people build new jobs and invest in the future.

It's a vision that I share with millions of Americans. And that's why I am supporting Bernie Sanders.

Type your input here: (Type exit or quit to quit the program)quit

## 2.2 facebook/opt-1.3b

### 2.2.1 Using the greedy search

```
[ ]: from transformers import AutoModelForCausalLM, AutoTokenizer

tokenizer = AutoTokenizer.from_pretrained('facebook/opt-1.3b')
model = AutoModelForCausalLM.from_pretrained('facebook/opt-1.3b').to('cuda')

inputs = tokenizer.encode("Hello, I would like to show some empathy and",
    ↪return_tensors="pt").to('cuda')

res = model.generate(inputs, max_length=512)
for num, item in enumerate(res):
    print('%d : %s' %(num, tokenizer.decode(item, skip_special_tokens=True)))
```

```
Downloading (âĖ)kenizer_config.json: 0%|          | 0.00/685 [00:00<?, ?B/s]
Downloading (âĖ)lve/main/config.json: 0%|         | 0.00/653 [00:00<?, ?B/s]
Downloading (âĖ)olve/main/vocab.json: 0%|        | 0.00/899k [00:00<?, ?B/s]
Downloading (âĖ)olve/main/merges.txt: 0%|         | 0.00/456k [00:00<?, ?B/s]
Downloading (âĖ)cial_tokens_map.json: 0%|         | 0.00/441 [00:00<?, ?B/s]
Downloading pytorch_model.bin: 0%|              | 0.00/2.63G [00:00<?, ?B/s]
Downloading (âĖ)neration_config.json: 0%|         | 0.00/137 [00:00<?, ?B/s]

0 : Hello, I would like to show some empathy and understanding to you. I am a
Christian and I am sorry that you are going through this. I hope that you will
find some peace and comfort in the Lord.
Thank you. I appreciate your kind words.
```

### 2.2.2 Using the beam search

```
[ ]: tokenizer = AutoTokenizer.from_pretrained('facebook/opt-1.3b')
model = AutoModelForCausalLM.from_pretrained('facebook/opt-1.3b').to('cuda')

inputs = tokenizer.encode("Hello, I would like to show some empathy and",
    ↪return_tensors="pt").to('cuda')

res = model.generate(inputs, max_length=512, num_beams=5,
    ↪num_return_sequences=5, early_stopping=True)
for num, item in enumerate(res):
    print('%d : %s' %(num, tokenizer.decode(item, skip_special_tokens=True)))
```

0 : Hello, I would like to show some empathy and support to you and your family.  
 1 : Hello, I would like to show some empathy and support to you and your family.  
 -  
 2 : Hello, I would like to show some empathy and support to you and your family.  
 -  
 3 : Hello, I would like to show some empathy and support to you and your family.  
 -  
 4 : Hello, I would like to show some empathy and support to you and your family.  
 -

### 2.2.3 using top-k sampling

```
[ ]: res = model.generate(inputs, max_length=512, do_sample = True, top_k = 0,
    ↪ num_return_sequences=5)
for num, item in enumerate(res):
    print('%d : %s' %(num, tokenizer.decode(item, skip_special_tokens=True)))
```

0 : Hello, I would like to show some empathy and hope that he recovers soon. I agree that a provider is needed that would be passionate about ethics and principled work in the industry as I think a community of providers need to rise on Imgur and become proactive in correcting the unethical content aimed at the moderators and admins by other users.  
 yes the value is important, but you also have to feel that the provider has their heart in it. you care about giving up you time, also your efforts to stay up to date on exigent issues that the community is concerned about. the difference between life and Death. if I could work in professional ethics and professionalism without having a heart\* I wouldn't be here. I don't think you can ;-/ \*Heart is relevant here because he works in a reproductive health medical venue so he has cared a setup more than most.  
 1 : Hello, I would like to show some empathy and compassion. Please read about internalizing your triggers then talk to a professional like me.  
 I will, thanks for the offer! I just rely on positive reinforcement and self compassion.  
 2 : Hello, I would like to show some empathy and break the cycle of abuse. I recently have been kicked out of my home due to severe railings. I really don't know what I'm going to do as none of my family would help in any way. I can't go out as you, everyone makes comments about me not putting the wood down. How do I break out of this? This happens every time it's seemed to me.  
 I tried the same advice /OP gave you on another post. I made certain shit clear to everyone that I could live without having to make a decision to take my life. And as you said they all said this. I've since luckily been able to move out thanks to all my friends and internet. I advice you do the same thing if you and incestual situation doesn't end soon. Contact supports at whatever school thats close to you for help. Good luck!  
 I've initiated contact with someone at college while in town to try to help me get my money or something. I'm getting bullied by very close friends though  
 3 : Hello, I would like to show some empathy and help if you are in need. My btc

address is 0x9Bro3HDQT6Bz9PbkVSX0Yys2cqPhCBK, I have both android and pc. You can check the addresses with your eyes and with the delve app.

I sent you my Bitcoin address, I have three devices. Can you look at one to see what is happening?

4 : Hello, I would like to show some empathy and love for you as I love meeting new people. This has provided me with hope during boring times at my job and to meet people over coffee or in a park when I'm out. I enjoy watching people vent and confide in me. Being a pup lover could provide me some fun times. asking for clarification on their jobs seems pointless and unhelpful. Please consider and write back if you can if you're interested.

Thanks for telling me it meant a lot to you, even if details are fuzzy and not 100% accurate. :)

## 2.2.4 using top-p sampling

```
[ ]: res = model.generate(inputs, max_length=512, do_sample = True, top_k = 0, top_p_
    ↳ = 0.7, num_return_sequences=5)
for num, item in enumerate(res):
    print('%d : %s' %(num, tokenizer.decode(item, skip_special_tokens=True)))
```

0 : Hello, I would like to show some empathy and help out. I'm a new player, and am curious as to how people make their games and the type of games they play. Do you have a preferred system?

I have a preference for what games I play. I have a basic rule set and that is how I make my games.

1 : Hello, I would like to show some empathy and compassion towards someone I do not know, and the reason I'm doing this is because I do not know the person.

This post is meant to start a conversation about how we treat people in the real world. I am aware that I have come off as insensitive, ignorant, and rude to many, and I hope that people can take a step back and see the reality of the situation.

I am writing this because I want to help people see that the way they respond to others in life does have a huge impact on the way we act in our personal lives.

People on the internet can be incredibly cruel and unforgiving. I am aware of this, and I have made many attempts to change my behavior to become more friendly and friendly-hearted.

However, there are certain situations in life that I cannot change. I cannot control how I am perceived by others, and I cannot change the way they perceive me.

There is a difference between what people perceive me to be and what I perceive myself to be.

So, if I make a joke about the passing of a friend, or talk about how I have been in the friendzone, it is only going to make things worse.

If I have a situation in my life where I have hurt someone, I can't do anything about it. I can only apologize, and hope that the person forgives me.

People don't see the person I am. They only see the parts that they want to see.

They only see the parts that are important to them, and they only see the parts that fit their own views.

People see the actions, and not the person.

The people who have done me harm will only see the parts of me that are important to them, and the person that has done me harm will only see the parts of me that are important to them.

This is a problem, and it is going to be a problem until people start to see the whole person.

So, please, let me start by apologizing for how I came off in this post. I am sorry that I hurt anyone, and I hope that you will forgive me.

I am sorry that I was insensitive and rude. I hope that you will forgive me.

I am sorry that I hurt you.

I hope that you will be understanding of my lack of empathy and compassion towards people.

2 : Hello, I would like to show some empathy and listen to you.

>hello Hello, it's you! >listen to you >caring You don't seem like you care. >empathy Please try again.

Oh, I don't think I cared, I was just hoping that you were actually listening.

I listen. I'm just not very good at it. >you're If you don't want to be called a he, use the pronoun you're not talking about. >I'm not very good at it If you don't want to be called a she, use the pronoun you are talking about.

Oh, I'm sorry, I didn't realize I was talking about me. >If you don't want to be called a he, use the pronoun you are talking about. I don't care what I am, you can call me whatever you want. >If you don't want to be called a she, use the pronoun you are talking about. I can't help it if I'm talking about me.

>I can't help it if I'm talking about me. If you're not, don't. I'm just trying to tell you that you're not talking about you. >If you're not, don't. I'm just trying to tell you that you're not talking about you. That's my point. If you don't want to be called a he, use the pronoun you are not talking about.

3 : Hello, I would like to show some empathy and understanding to a man who's wife has cancer. He's in a tough situation and has already been through a

divorce. I'm sure you don't want to be alone, so I would appreciate it if you could reply and let me know if I can help.  
I just posted a reply in your comment, but I would appreciate it if you could send it to him and I will let you know how I can help. Thank you!  
I got it! Thank you!  
4 : Hello, I would like to show some empathy and sympathy for your losses, and I hope your life can get better soon. I will say that some things do seem unfair. Thank you.

## 3 Task 3

### 3.1 Load the data set

```
[ ]: # mount the drive
from google.colab import drive
drive.mount('gdrive')
```

Mounted at gdrive

```
[ ]: # change the drive to my working directory
%cd gdrive/MyDrive/CSC583
```

/content/gdrive/MyDrive/CSC583

```
[ ]: # helper function
def readPlainDoc(filename_s):
    """Function that returns a string in lower case after reading the contents of
    the file"""
    fileNamePath = "data/" + filename_s
    with open(fileNamePath, 'r') as sampleFile:
        fileContents = sampleFile.read()
    # normalize all words to lower case
    return fileContents.lower()
```

```
[ ]: # read the file
warAndPeace_text = readPlainDoc("war-and-peace.txt")
#warAndPeace_text[100:]
```

```
[ ]: # Check if the GPU has been enabled.
print(torch.__version__, torch.cuda.is_available())
```

2.0.1+cu118 True

## 3.2 gpt2-large

```
[ ]: from transformers import GPT2LMHeadModel, GPT2TokenizerFast
```

```
device = 'cuda'
model_id = 'gpt2-large'
model = GPT2LMHeadModel.from_pretrained(model_id).to(device)
tokenizer = GPT2TokenizerFast.from_pretrained(model_id)
```

```
Downloading (âĤ)lve/main/config.json: 0%|          | 0.00/666 [00:00<?, ?B/s]
Downloading pytorch_model.bin: 0%|          | 0.00/3.25G [00:00<?, ?B/s]
Downloading (âĤ)neration_config.json: 0%|          | 0.00/124 [00:00<?, ?B/s]
Downloading (âĤ)olve/main/vocab.json: 0%|          | 0.00/1.04M [00:00<?, ?B/s]
Downloading (âĤ)olve/main/merges.txt: 0%|          | 0.00/456k [00:00<?, ?B/s]
Downloading (âĤ)/main/tokenizer.json: 0%|          | 0.00/1.36M [00:00<?, ?B/s]
```

```
[ ]: encodings = tokenizer(warAndPeace_text[1000:], return_tensors="pt")
```

Token indices sequence length is longer than the specified maximum sequence length for this model (813490 > 1024). Running this sequence through the model will result in indexing errors

### 3.2.1 using the size of stride of 512

```
[ ]: import torch
from tqdm import tqdm

max_length = model.config.n_positions
stride = 512
seq_len = encodings.input_ids.size(1)

nlls = []
prev_end_loc = 0
for begin_loc in tqdm(range(0, seq_len, stride)):
    end_loc = min(begin_loc + max_length, seq_len)
    trg_len = end_loc - prev_end_loc # may be different from stride on last loop
    input_ids = encodings.input_ids[:, begin_loc:end_loc].to(device)
    target_ids = input_ids.clone()
    target_ids[:, :-trg_len] = -100

    with torch.no_grad():
        outputs = model(input_ids, labels=target_ids)
```





```

        # N.B. the model only calculates loss over trg_len - 1 labels, because
        ↪ it internally shifts the labels
        # to the left by 1.
        neg_log_likelihood = outputs.loss

        nlls.append(neg_log_likelihood)

        prev_end_loc = end_loc
        if end_loc == seq_len:
            break

ppl = torch.exp(torch.stack(nlls).mean())
print('\n')
print('The perplexity is %.3f.' %ppl.item())

```

100%|aũLaũLaũLaũLaũLaũLaũLaũLaũLaũL| 794/795 [08:03<00:00, 1.64it/s]

The perplexity is 26.106.

### 3.2.3 using the size of stride of 256

```

[ ]: import torch
      from tqdm import tqdm

max_length = model.config.n_positions
stride = 256
seq_len = encodings.input_ids.size(1)

nlls = []
prev_end_loc = 0
for begin_loc in tqdm(range(0, seq_len, stride)):
    end_loc = min(begin_loc + max_length, seq_len)
    trg_len = end_loc - prev_end_loc # may be different from stride on last loop
    input_ids = encodings.input_ids[:, begin_loc:end_loc].to(device)
    target_ids = input_ids.clone()
    target_ids[:, :-trg_len] = -100

    with torch.no_grad():
        outputs = model(input_ids, labels=target_ids)

        # loss is calculated using CrossEntropyLoss which averages over valid
        ↪ labels

```



```

        # N.B. the model only calculates loss over trg_len - 1 labels, because
        → it internally shifts the labels
        # to the left by 1.
        neg_log_likelihood = outputs.loss

        nlls.append(neg_log_likelihood)

        prev_end_loc = end_loc
        if end_loc == seq_len:
            break

    ppl = torch.exp(torch.stack(nlls).mean())
    print('\n')
    print('The perplexity is %.3f.' %ppl.item())

```

100%|aũLaũLaũLaũLaũLaũLaũLaũLaũLaũL| 529/530 [05:20<00:00, 1.65it/s]

The perplexity is 25.955.

### 3.3 gpt2-medium

```

[ ]: from transformers import GPT2LMHeadModel, GPT2TokenizerFast

device = 'cuda'
model_id = 'gpt2-medium'
model = GPT2LMHeadModel.from_pretrained(model_id).to(device)
tokenizer = GPT2TokenizerFast.from_pretrained(model_id)

```

#### 3.3.1 using the size of stride of 512

```

[ ]: import torch
from tqdm import tqdm

encodings = tokenizer(warAndPeace_text[750:], return_tensors="pt")
max_length = model.config.n_positions
stride = 512
seq_len = encodings.input_ids.size(1)

nlls = []
prev_end_loc = 0
for begin_loc in tqdm(range(0, seq_len, stride)):
    end_loc = min(begin_loc + max_length, seq_len)

```





