03 - Baseline Summarization using Transformer model

In this notebook file, we will do a transformer baseline model:

- read data from mini evaluation set (10% random sampling)
- extract abstract text (first paragraph)
- do summarization with model "Falconsai/medical_summarization"

In [1]:

```
!pip install pandas
!pip install transformers
```

```
Requirement already satisfied: pandas in c:\users\minh ubc\miniconda3\lib\site-packa
ges (2.2.1)
Requirement already satisfied: numpy<2,>=1.23.2 in c:\users\minh ubc\miniconda3\lib
\site-packages (from pandas) (1.25.2)
Requirement already satisfied: python-dateutil>=2.8.2 in c:\users\minh ubc\miniconda
3\lib\site-packages (from pandas) (2.9.0)
Requirement already satisfied: pytz>=2020.1 in c:\users\minh ubc\miniconda3\lib\site
-packages (from pandas) (2024.1)
Requirement already satisfied: tzdata>=2022.7 in c:\users\minh ubc\miniconda3\lib\si
te-packages (from pandas) (2024.1)
Requirement already satisfied: six>=1.5 in c:\users\minh ubc\miniconda3\lib\site-pac
kages (from python-dateutil>=2.8.2->pandas) (1.16.0)
Requirement already satisfied: transformers in c:\users\minh ubc\miniconda3\lib\site
-packages (4.38.1)
Requirement already satisfied: filelock in c:\users\minh ubc\miniconda3\lib\site-pac
kages (from transformers) (3.13.1)
Requirement already satisfied: huggingface-hub<1.0,>=0.19.3 in c:\users\minh ubc\min
iconda3\lib\site-packages (from transformers) (0.19.4)
Requirement already satisfied: numpy>=1.17 in c:\users\minh ubc\miniconda3\lib\site-
packages (from transformers) (1.25.2)
Requirement already satisfied: packaging>=20.0 in c:\users\minh ubc\miniconda3\lib\s
ite-packages (from transformers) (24.0)
Requirement already satisfied: pyyaml>=5.1 in c:\users\minh ubc\miniconda3\lib\site-
packages (from transformers) (6.0.1)
Requirement already satisfied: regex!=2019.12.17 in c:\users\minh ubc\miniconda3\lib
\site-packages (from transformers) (2023.10.3)
Requirement already satisfied: requests in c:\users\minh ubc\miniconda3\lib\site-pac
kages (from transformers) (2.31.0)
Requirement already satisfied: tokenizers<0.19,>=0.14 in c:\users\minh ubc\miniconda
3\lib\site-packages (from transformers) (0.15.2)
Requirement already satisfied: safetensors>=0.4.1 in c:\users\minh ubc\miniconda3\li
b\site-packages (from transformers) (0.4.1)
Requirement already satisfied: tqdm>=4.27 in c:\users\minh ubc\miniconda3\lib\site-p
ackages (from transformers) (4.66.1)
Requirement already satisfied: fsspec>=2023.5.0 in c:\users\minh ubc\miniconda3\lib
\site-packages (from huggingface-hub<1.0,>=0.19.3->transformers) (2023.10.0)
Requirement already satisfied: typing-extensions>=3.7.4.3 in c:\users\minh ubc\minic
onda3\lib\site-packages (from huggingface-hub<1.0,>=0.19.3->transformers) (4.10.0)
Requirement already satisfied: colorama in c:\users\minh ubc\miniconda3\lib\site-pac
kages (from tqdm>=4.27->transformers) (0.4.6)
Requirement already satisfied: charset-normalizer<4,>=2 in c:\users\minh ubc\minicon
da3\lib\site-packages (from requests->transformers) (3.3.2)
Requirement already satisfied: idna<4,>=2.5 in c:\users\minh ubc\miniconda3\lib\site
-packages (from requests->transformers) (3.6)
Requirement already satisfied: urllib3<3,>=1.21.1 in c:\users\minh ubc\miniconda3\li
b\site-packages (from requests->transformers) (2.2.1)
Requirement already satisfied: certifi>=2017.4.17 in c:\users\minh ubc\miniconda3\li
b\site-packages (from requests->transformers) (2024.2.2)
```

In [2]:

```python
import pandas as pd
from transformers import pipeline
import torch
```

In [3]:

```python
# checking if GPU is available
print(torch.cuda.is_available())
print(torch.cuda.current_device())
```

```
True
0
```

In [4]:
```
# set up summarization model
model = pipeline("summarization",
                 model="Falconsai/medical_summarization",
                 device=0 # run on GPU
                )
```

In [5]:
```
s = 'The evolutionary origins of the hypoxia-sensitive cells that trigger amniote r
s
```

Out[5]:
```
'The evolutionary origins of the hypoxia-sensitive cells that trigger amniote resp
iratory reflexes – carotid body glomus cells , and ‘pulmonary neuroendocrine cell
s’ ( PNECs ) - are obscure . Homology has been proposed between glomus cells , whi
ch are neural crest-derived , and the hypoxia-sensitive ‘neuroepithelial cells’ (
NECs ) of fish gills , whose embryonic origin is unknown . NECs have also been lik
ened to PNECs , which differentiate in situ within lung airway epithelia . Using g
enetic lineage-tracing and neural crest-deficient mutants in zebrafish , and physi
cal fate-mapping in frog and lamprey , we find that NECs are not neural crest-deri
ved , but endoderm-derived , like PNECs , whose endodermal origin we confirm . We
discover neural crest-derived catecholaminergic cells associated with zebrafish ph
aryngeal arch blood vessels , and propose a new model for amniote hypoxia-sensitiv
e cell evolution: endoderm-derived NECs were retained as PNECs , while the carotid
body evolved via the aggregation of neural crest-derived catecholaminergic ( chrom
affin ) cells already associated with blood vessels in anamniote pharyngeal arches
. '
```

In [6]:
```
result = model(s,
               max_length=500,
               min_length=100
              )
summ = result[0]["summary_text"]
print(len(summ))
print(summ)
```

```
Your max_length is set to 500, but your input_length is only 354. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=177)
783
the evolutionary origins of the hypoxia-sensitive cells that trigger amniote respira
tory reflexes – carotid body glomus cells , and ‘pulmonary neuroendocrine cells’ ( P
NECs ) - are obscure . we find that endoderm-derived neuroepithelial cells ( nEC ) a
re not neural crest , but endodermal , like PNCEs , whose endodermale origin we conf
irm . here we propose a new model for amnium-sensitive cell evolution : endodermm-de
rived catecholaminergic cells associated with zebrafish pharyngeal arch blood vessel
s , while the aggregation of chromaffin ) cells already associated with blood vessel
s in anamniotic arches has been proposed between the gills . ncs have also been like
ned to nephrine , which differentiate in situ within lung airway epithelia . the phe
notypic origin is unknown .
```

In [7]:
```
# put in a function
def summarize(text, model=model, min_length=100, max_length=500):
    """
```

```
        summarize a text using a transformer model,
        with min_length and max_length are number of tokens limits for the output
    """
    doc = model(text,
                max_length=max_length,
                min_length=min_length
               )
    summ = doc[0]["summary_text"]
    return summ

summarize(s)
```

```
Your max_length is set to 500, but your input_length is only 354. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=177)
```

Out[7]:

```
'the evolutionary origins of the hypoxia-sensitive cells that trigger amniote resp
iratory reflexes – carotid body glomus cells , and ‘pulmonary neuroendocrine cell
s’ ( PNECs ) - are obscure . we find that endoderm-derived neuroepithelial cells (
nEC ) are not neural crest , but endodermal , like PNCEs , whose endodermale origi
n we confirm . here we propose a new model for amnium-sensitive cell evolution : e
ndodermm-derived catecholaminergic cells associated with zebrafish pharyngeal arch
blood vessels , while the aggregation of chromaffin ) cells already associated wit
h blood vessels in anamniotic arches has been proposed between the gills . ncs hav
e also been likened to nephrine , which differentiate in situ within lung airway e
pithelia . the phenotypic origin is unknown .'
```

In [8]:

```
# Load data
filepath = "../data/mini_dataset/"
filename = "eLife_val_mini_milestone3.jsonl"
df = pd.read_json(filepath + filename,
                  orient="records",
                  lines=True
                 )
print(len(df))
df.head()
```

Out[8]:

	lay_summary	article	headings	keywords	id
0	It can take several months , or even years , f...	Mature neural networks synchronize and integra...	[Abstract, Introduction, Results, Discussion, ...	[neuroscience]	elife-69011-v2
1	Many of our decisions are made on the basis of...	Many decisions are thought to arise via the ac...	[Abstract, Introduction, Results, Discussion, ...	[neuroscience]	elife-17688-v1
2	Oculo-Cerebro-Renal syndrome of Lowe (Lowe sy...	Mutations in the inositol 5-phosphatase OCRL c...	[Abstract, Introduction, Results, Discussion, ...	[cell biology]	elife-02975-v2
3	When an embryo develops , its cells must work ...	Gradients of signaling proteins are essential ...	[Abstract, Introduction, Results, Discussion, ...	[developmental biology]	elife-38137-v3
4	Our genomes contain a record of historical eve...	Similarity between two individuals in the comb...	[Abstract, Introduction, Results, Discussion, ...	[evolutionary biology, genetics and genomics]	elife-15266-v1

In [9]:

```
# extract abstract (first paragraph)
df["abstract"] = df["article"].apply(lambda text: text.split("\n")[0])
print(df["abstract"].iloc[3])
df.head()
```

```
Gradients of signaling proteins are essential for inducing tissue morphogenesis . Ho
wever , mechanisms of gradient formation remain controversial . Here we characterize
d the distribution of fluorescently-tagged signaling proteins , FGF and FGFR , expre
ssed at physiological levels from the genomic knock-in alleles in Drosophila . FGF p
roduced in the larval wing imaginal-disc moves to the air-sac-primordium ( ASP ) thr
ough FGFR-containing cytonemes that extend from the ASP to contact the wing-disc sou
rce . The number of FGF-receiving cytonemes extended by ASP cells decreases graduall
y with increasing distance from the source , generating a recipient-specific FGF gra
dient . Acting as a morphogen in the ASP , FGF activates concentration-dependent gen
e expression , inducing pointed-P1 at higher and cut at lower levels . The transcrip
tion-factors Pointed-P1 and Cut antagonize each other and differentially regulate fo
rmation of FGFR-containing cytonemes , creating regions with higher-to-lower numbers
of FGF-receiving cytonemes . These results reveal a robust mechanism where morphogen
s self-generate precise tissue-specific gradient contours through feedback regulatio
n of cytoneme-mediated dispersion .
```

Out[9]:

	lay_summary	article	headings	keywords	id	abstract
0	It can take several months , or even years , f...	Mature neural networks synchronize and integra...	[Abstract, Introduction, Results, Discussion, ...	[neuroscience]	elife-69011-v2	Mature neural networks synchronize and integra...
1	Many of our decisions are made on the basis of...	Many decisions are thought to arise via the ac...	[Abstract, Introduction, Results, Discussion, ...	[neuroscience]	elife-17688-v1	Many decisions are thought to arise via the ac...
2	Oculo-Cerebro-Renal syndrome of Lowe (Lowe sy...	Mutations in the inositol 5-phosphatase OCRL c...	[Abstract, Introduction, Results, Discussion, ...	[cell biology]	elife-02975-v2	Mutations in the inositol 5-phosphatase OCRL c...
3	When an embryo develops , its cells must work ...	Gradients of signaling proteins are essential ...	[Abstract, Introduction, Results, Discussion, ...	[developmental biology]	elife-38137-v3	Gradients of signaling proteins are essential ...
4	Our genomes contain a record of historical eve...	Similarity between two individuals in the comb...	[Abstract, Introduction, Results, Discussion, ...	[evolutionary biology, genetics and genomics]	elife-15266-v1	Similarity between two individuals in the comb...

In [10]:

```
# apply summarization
df["baseline_summary"] = df["abstract"].apply(lambda text: summarize(text))
df.head()
```

```
Your max_length is set to 500, but your input_length is only 226. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=113)
Your max_length is set to 500, but your input_length is only 206. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=103)
Your max_length is set to 500, but your input_length is only 318. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=159)
Your max_length is set to 500, but your input_length is only 321. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=160)
Your max_length is set to 500, but your input_length is only 237. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=118)
Your max_length is set to 500, but your input_length is only 296. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=148)
Your max_length is set to 500, but your input_length is only 369. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=184)
Your max_length is set to 500, but your input_length is only 317. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=158)
C:\Users\Minh UBC\miniconda3\Lib\site-packages\transformers\pipelines\base.py:1157:
UserWarning: You seem to be using the pipelines sequentially on GPU. In order to max
imize efficiency please use a dataset
  warnings.warn(
Your max_length is set to 500, but your input_length is only 287. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=143)
Your max_length is set to 500, but your input_length is only 281. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=140)
Your max_length is set to 500, but your input_length is only 244. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=122)
Your max_length is set to 500, but your input_length is only 249. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=124)
Your max_length is set to 500, but your input_length is only 213. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=106)
Your max_length is set to 500, but your input_length is only 229. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=114)
Your max_length is set to 500, but your input_length is only 246. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=123)
Your max_length is set to 500, but your input_length is only 246. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=123)
Your max_length is set to 500, but your input_length is only 344. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=172)
Your max_length is set to 500, but your input_length is only 306. Since this is a su
```

```
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=153)
Your max_length is set to 500, but your input_length is only 266. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=133)
Your max_length is set to 500, but your input_length is only 227. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=113)
Your max_length is set to 500, but your input_length is only 250. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=125)
Your max_length is set to 500, but your input_length is only 286. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=143)
Your max_length is set to 500, but your input_length is only 293. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=146)
Your max_length is set to 500, but your input_length is only 256. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=128)
```

Out[10]:

	lay_summary	article	headings	keywords	id	abstract	baseline_su
0	It can take several months , or even years , f...	Mature neural networks synchronize and integra...	[Abstract, Introduction, Results, Discussion, ...	[neuroscience]	elife-69011-v2	Mature neural networks synchronize and integra...	we investi progre large
1	Many of our decisions are made on the basis of...	Many decisions are thought to arise via the ac...	[Abstract, Introduction, Results, Discussion, ...	[neuroscience]	elife-17688-v1	Many decisions are thought to arise via the ac...	a novel mani mimics the
2	Oculo-Cerebro-Renal syndrome of Lowe (Lowe sy...	Mutations in the inositol 5-phosphatase OCRL c...	[Abstract, Introduction, Results, Discussion, ...	[cell biology]	elife-02975-v2	Mutations in the inositol 5-phosphatase OCRL c...	the in phosphata is recr
3	When an embryo develops , its cells must work ...	Gradients of signaling proteins are essential ...	[Abstract, Introduction, Results, Discussion, ...	[developmental biology]	elife-38137-v3	Gradients of signaling proteins are essential ...	gradient fo controvers
4	Our genomes contain a record of historical eve...	Similarity between two individuals in the comb...	[Abstract, Introduction, Results, Discussion, ...	[evolutionary biology, genetics and genomics]	elife-15266-v1	Similarity between two individuals in the comb...	back similarity two inc

In [11]:

```
# write to output
output_path = "../data/milestone3/transformer_baseline/"
output_file = "elife.csv"
df.to_csv(output_path+output_file,
          index=False,
         )
print("Output file completed")
```

```
Output file completed
```

In [12]:

```
# write to txt file
output_file_txt = "elife.txt"

# write the baseline_summary column to txt file
txt_df = df['baseline_summary']
txt_df.to_csv(output_path+output_file_txt,
              index=False,
              header=False,
              sep="\n"
             )
print("Output file completed")
```

```
Output file completed
```

In [13]:

```
# repeat for PLOS dev set
# load data
filepath = "../data/mini_dataset/"
filename = "PLOS_val_mini_milestone3.jsonl"
df = pd.read_json(filepath + filename,
                  orient="records",
                  lines=True
                 )
print(len(df))
df.head()
```

```
138
```

Out[13]:

	lay_summary	article	headings	keywords	id
0	Yersinia pestis , the bacterial agent of plagu...	Fleas can transmit Yersinia pestis by two mech...	[Abstract, Introduction, Results, Discussion, ...	[united states, invertebrates, medicine and he...	journal.ppat.1006859
1	The genome of all vertebrates is heavily colon...	Endogenous retroviruses (ERVs) are remnants ...	[Abstract, Introduction, Results, Discussion, ...	[viruses, sheep, virology]	journal.ppat.0030170
2	The molecular mechanisms underlying directed c...	The Drosophila embryonic gonad is assembled fr...	[Abstract, Introduction, Results, Discussion, ...	[]	journal.pgen.1003720
3	Contrary to the long-standing belief that no n...	Recently , we presented a study of adult neuro...	[Abstract, Introduction, Model, Results, Discu...	[computational biology/computational neuroscie...	journal.pcbi.1001063
4	Embryonic stem cells have two remarkable prope...	Understanding the transcriptional regulation o...	[Abstract, Introduction, Results, Discussion, ...	[developmental biology, cell biology, mammals,...	journal.pgen.0030145

In [14]:

```
# extract abstract (first paragraph)
df["abstract"] = df["article"].apply(lambda text: text.split("\n")[0])
print(df["abstract"].iloc[3])
df.head()
```

```
Recently , we presented a study of adult neurogenesis in a simplified hippocampal me
mory model . The network was required to encode and decode memory patterns despite c
hanging input statistics . We showed that additive neurogenesis was a more effective
adaptation strategy compared to neuronal turnover and conventional synaptic plastici
ty as it allowed the network to respond to changes in the input statistics while pre
serving representations of earlier environments . Here we extend our model to includ
e realistic , spatially driven input firing patterns in the form of grid cells in th
e entorhinal cortex . We compare network performance across a sequence of spatial en
vironments using three distinct adaptation strategies: conventional synaptic plastic
ity , where the network is of fixed size but the connectivity is plastic; neuronal t
urnover , where the network is of fixed size but units in the network may die and be
replaced; and additive neurogenesis , where the network starts out with fewer initia
l units but grows over time . We confirm that additive neurogenesis is a superior ad
aptation strategy when using realistic , spatially structured input patterns . We th
en show that a more biologically plausible neurogenesis rule that incorporates cell 
death and enhanced plasticity of new granule cells has an overall performance signif
icantly better than any one of the three individual strategies operating alone . Thi
s adaptation rule can be tailored to maximise performance of the network when operat
ing as either a short- or long-term memory store . We also examine the time course o
f adult neurogenesis over the lifetime of an animal raised under different hypotheti
cal rearing conditions . These growth profiles have several distinct features that f
orm a theoretical prediction that could be tested experimentally . Finally , we show
that place cells can emerge and refine in a realistic manner in our model as a direc
t result of the sparsification performed by the dentate gyrus layer .
```

Out[14]:

	lay_summary	article	headings	keywords	id	
0	Yersinia pestis , the bacterial agent of plagu...	Fleas can transmit Yersinia pestis by two mech...	[Abstract, Introduction, Results, Discussion, ...	[united states, invertebrates, medicine and he...	journal.ppat.1006859	Y b
1	The genome of all vertebrates is heavily colon...	Endogenous retroviruses (ERVs) are remnants ...	[Abstract, Introduction, Results, Discussion, ...	[viruses, sheep, virology]	journal.ppat.0030170	
2	The molecular mechanisms underlying directed c...	The Drosophila embryonic gonad is assembled fr...	[Abstract, Introduction, Results, Discussion, ...	[]	journal.pgen.1003720	as
3	Contrary to the long-standing belief that no n...	Recently , we presented a study of adult neuro...	[Abstract, Introduction, Model, Results, Discu...	[computational biology/computational neuroscie...	journal.pcbi.1001063	s
4	Embryonic stem cells have two remarkable prope...	Understanding the transcriptional regulation o...	[Abstract, Introduction, Results, Discussion, ...	[developmental biology, cell biology, mammals,...	journal.pgen.0030145	Ur tr r

In [15]:

```
# apply summarization
df["baseline_summary"] = df["abstract"].apply(lambda text: summarize(text))
df.head()
```

```
C:\Users\Minh UBC\miniconda3\Lib\site-packages\transformers\pipelines\base.py:1157:
UserWarning: You seem to be using the pipelines sequentially on GPU. In order to max
imize efficiency please use a dataset
  warnings.warn(
Token indices sequence length is longer than the specified maximum sequence length f
or this model (531 > 512). Running this sequence through the model will result in in
dexing errors
Your max_length is set to 500, but your input_length is only 461. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=230)
Your max_length is set to 500, but your input_length is only 398. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=199)
Your max_length is set to 500, but your input_length is only 359. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=179)
Your max_length is set to 500, but your input_length is only 417. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=208)
Your max_length is set to 500, but your input_length is only 239. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=119)
Your max_length is set to 500, but your input_length is only 267. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=133)
Your max_length is set to 500, but your input_length is only 371. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=185)
Your max_length is set to 500, but your input_length is only 345. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=172)
Your max_length is set to 500, but your input_length is only 392. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=196)
Your max_length is set to 500, but your input_length is only 449. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=224)
Your max_length is set to 500, but your input_length is only 285. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=142)
Your max_length is set to 500, but your input_length is only 497. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=248)
Your max_length is set to 500, but your input_length is only 361. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=180)
Your max_length is set to 500, but your input_length is only 358. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=179)
Your max_length is set to 500, but your input_length is only 414. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=207)
Your max_length is set to 500, but your input_length is only 269. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=134)
Your max_length is set to 500, but your input_length is only 229. Since this is a su
```

```
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=114)
Your max_length is set to 500, but your input_length is only 423. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=211)
Your max_length is set to 500, but your input_length is only 331. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=165)
Your max_length is set to 500, but your input_length is only 413. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=206)
Your max_length is set to 500, but your input_length is only 452. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=226)
Your max_length is set to 500, but your input_length is only 375. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=187)
Your max_length is set to 500, but your input_length is only 434. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=217)
Your max_length is set to 500, but your input_length is only 425. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=212)
Your max_length is set to 500, but your input_length is only 291. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=145)
Your max_length is set to 500, but your input_length is only 391. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=195)
Your max_length is set to 500, but your input_length is only 464. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=232)
Your max_length is set to 500, but your input_length is only 393. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=196)
Your max_length is set to 500, but your input_length is only 410. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=205)
Your max_length is set to 500, but your input_length is only 243. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=121)
Your max_length is set to 500, but your input_length is only 300. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=150)
Your max_length is set to 500, but your input_length is only 317. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=158)
Your max_length is set to 500, but your input_length is only 345. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=172)
Your max_length is set to 500, but your input_length is only 358. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=179)
Your max_length is set to 500, but your input_length is only 477. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=238)
```

```
Your max_length is set to 500, but your input_length is only 475. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=237)
Your max_length is set to 500, but your input_length is only 244. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=122)
Your max_length is set to 500, but your input_length is only 431. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=215)
Your max_length is set to 500, but your input_length is only 426. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=213)
Your max_length is set to 500, but your input_length is only 216. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=108)
Your max_length is set to 500, but your input_length is only 358. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=179)
Your max_length is set to 500, but your input_length is only 358. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=179)
Your max_length is set to 500, but your input_length is only 449. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=224)
Your max_length is set to 500, but your input_length is only 401. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=200)
Your max_length is set to 500, but your input_length is only 368. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=184)
Your max_length is set to 500, but your input_length is only 443. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=221)
Your max_length is set to 500, but your input_length is only 412. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=206)
Your max_length is set to 500, but your input_length is only 363. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=181)
Your max_length is set to 500, but your input_length is only 308. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=154)
Your max_length is set to 500, but your input_length is only 432. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=216)
Your max_length is set to 500, but your input_length is only 372. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=186)
Your max_length is set to 500, but your input_length is only 449. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=224)
Your max_length is set to 500, but your input_length is only 321. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=160)
Your max_length is set to 500, but your input_length is only 298. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
```

```
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=149)
Your max_length is set to 500, but your input_length is only 375. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=187)
Your max_length is set to 500, but your input_length is only 331. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=165)
Your max_length is set to 500, but your input_length is only 369. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=184)
Your max_length is set to 500, but your input_length is only 395. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=197)
Your max_length is set to 500, but your input_length is only 382. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=191)
Your max_length is set to 500, but your input_length is only 463. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=231)
Your max_length is set to 500, but your input_length is only 354. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=177)
Your max_length is set to 500, but your input_length is only 423. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=211)
Your max_length is set to 500, but your input_length is only 369. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=184)
Your max_length is set to 500, but your input_length is only 424. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=212)
Your max_length is set to 500, but your input_length is only 256. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=128)
Your max_length is set to 500, but your input_length is only 372. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=186)
Your max_length is set to 500, but your input_length is only 393. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=196)
Your max_length is set to 500, but your input_length is only 366. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=183)
Your max_length is set to 500, but your input_length is only 404. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=202)
Your max_length is set to 500, but your input_length is only 308. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=154)
Your max_length is set to 500, but your input_length is only 393. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=196)
Your max_length is set to 500, but your input_length is only 356. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=178)
Your max_length is set to 500, but your input_length is only 279. Since this is a su
```

```
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=139)
Your max_length is set to 500, but your input_length is only 265. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=132)
Your max_length is set to 500, but your input_length is only 482. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=241)
Your max_length is set to 500, but your input_length is only 343. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=171)
Your max_length is set to 500, but your input_length is only 468. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=234)
Your max_length is set to 500, but your input_length is only 415. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=207)
Your max_length is set to 500, but your input_length is only 391. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=195)
Your max_length is set to 500, but your input_length is only 442. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=221)
Your max_length is set to 500, but your input_length is only 437. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=218)
Your max_length is set to 500, but your input_length is only 484. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=242)
Your max_length is set to 500, but your input_length is only 481. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=240)
Your max_length is set to 500, but your input_length is only 208. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=104)
Your max_length is set to 500, but your input_length is only 400. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=200)
Your max_length is set to 500, but your input_length is only 409. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=204)
Your max_length is set to 500, but your input_length is only 358. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=179)
Your max_length is set to 500, but your input_length is only 365. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=182)
Your max_length is set to 500, but your input_length is only 282. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=141)
Your max_length is set to 500, but your input_length is only 498. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=249)
Your max_length is set to 500, but your input_length is only 361. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=180)
```

```
Your max_length is set to 500, but your input_length is only 356. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=178)
Your max_length is set to 500, but your input_length is only 303. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=151)
Your max_length is set to 500, but your input_length is only 463. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=231)
Your max_length is set to 500, but your input_length is only 294. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=147)
Your max_length is set to 500, but your input_length is only 385. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=192)
Your max_length is set to 500, but your input_length is only 348. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=174)
Your max_length is set to 500, but your input_length is only 366. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=183)
Your max_length is set to 500, but your input_length is only 445. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=222)
Your max_length is set to 500, but your input_length is only 461. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=230)
Your max_length is set to 500, but your input_length is only 412. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=206)
Your max_length is set to 500, but your input_length is only 423. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=211)
```

Out[15]:

	lay_summary	article	headings	keywords	id	
0	Yersinia pestis , the bacterial agent of plagu...	Fleas can transmit Yersinia pestis by two mech...	[Abstract, Introduction, Results, Discussion, ...	[united states, invertebrates, medicine and he...	journal.ppat.1006859	Y by
1	The genome of all vertebrates is heavily colon...	Endogenous retroviruses (ERVs) are remnants ...	[Abstract, Introduction, Results, Discussion, ...	[viruses, sheep, virology]	journal.ppat.0030170	ı
2	The molecular mechanisms underlying directed c...	The Drosophila embryonic gonad is assembled fr...	[Abstract, Introduction, Results, Discussion, ...	[]	journal.pgen.1003720	as
3	Contrary to the long-standing belief that no n...	Recently , we presented a study of adult neuro...	[Abstract, Introduction, Model, Results, Discu...	[computational biology/computational neuroscie...	journal.pcbi.1001063	l s
4	Embryonic stem cells have two remarkable prope...	Understanding the transcriptional regulation o...	[Abstract, Introduction, Results, Discussion, ...	[developmental biology, cell biology, mammals,...	journal.pgen.0030145	Ur tr r

In [16]:

```
# write to output
output_path = "../data/milestone3/transformer_baseline/"
output_file = "plos_mini.csv"
df.to_csv(output_path+output_file,
          index=False,
         )
print("Output file completed")
```

```
Output file completed
```

In [17]:

```
# write to txt file
output_file_txt = "plos.txt"

# write the baseline_summary column to txt file
txt_df = df['baseline_summary']
txt_df.to_csv(output_path+output_file_txt,
              index=False,
              header=False,
              sep="\n"
             )
print("Output file completed")
```

```
Output file completed
```

In [21]:

```python
# apply to test sets
filepath = "../data/biolaysumm2024_data/"
filenames = ["eLife_test.jsonl",
             "PLOS_test.jsonl"
            ]
output_path = "../data/milestone3/transformer_baseline/test_set/"

output_csv_filenames = ["elife_test.csv",
                        "plos_test.csv"
                       ]

output_txt_filenames = ["elife.txt",
                        "plos.txt"
                       ]



for i, fname in enumerate(filenames):
    print("Loading file = ", fname)
    df = pd.read_json(filepath + fname,
                      orient="records",
                      lines=True
                     )
    print("n rows =", len(df))

    print("Extracting abstract...")
    df["abstract"] = df["article"].apply(lambda text: text.split("\n")[0])

    print("Making summaries....")
    df["baseline_summary"] = df["abstract"].apply(lambda text: summarize(text))

    print("Writing csv output...")
    df.to_csv(output_path+output_csv_filenames[i],
              index=False,
             )
    print("Output csv file completed")

    txt_df = df['baseline_summary']
    txt_df.to_csv(output_path+output_txt_filenames[i],
                  index=False,
                  header=False,
                  sep="\n"
                 )
    print("Output txt file completed")

print("---- All completed----")
```

```
C:\Users\Minh UBC\miniconda3\Lib\site-packages\transformers\pipelines\base.py:1157:
UserWarning: You seem to be using the pipelines sequentially on GPU. In order to max
imize efficiency please use a dataset
  warnings.warn(
Your max_length is set to 500, but your input_length is only 394. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=197)
```

```
Loading file =  eLife_test.jsonl
n rows = 142
Extracting abstract...
Making summaries....
```

```
Your max_length is set to 500, but your input_length is only 259. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=129)
Your max_length is set to 500, but your input_length is only 352. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=176)
Your max_length is set to 500, but your input_length is only 392. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=196)
Your max_length is set to 500, but your input_length is only 308. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=154)
Your max_length is set to 500, but your input_length is only 491. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=245)
Your max_length is set to 500, but your input_length is only 480. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=240)
Your max_length is set to 500, but your input_length is only 382. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=191)
Your max_length is set to 500, but your input_length is only 325. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=162)
Your max_length is set to 500, but your input_length is only 378. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=189)
Your max_length is set to 500, but your input_length is only 341. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=170)
Your max_length is set to 500, but your input_length is only 380. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=190)
Your max_length is set to 500, but your input_length is only 307. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=153)
Your max_length is set to 500, but your input_length is only 439. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=219)
Your max_length is set to 500, but your input_length is only 349. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=174)
Your max_length is set to 500, but your input_length is only 404. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=202)
Your max_length is set to 500, but your input_length is only 215. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=107)
Your max_length is set to 500, but your input_length is only 280. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=140)
Your max_length is set to 500, but your input_length is only 344. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=172)
Your max_length is set to 500, but your input_length is only 358. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
```

```
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=179)
Your max_length is set to 500, but your input_length is only 270. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=135)
Your max_length is set to 500, but your input_length is only 330. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=165)
Your max_length is set to 500, but your input_length is only 223. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=111)
Your max_length is set to 500, but your input_length is only 328. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=164)
Your max_length is set to 500, but your input_length is only 327. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=163)
Your max_length is set to 500, but your input_length is only 284. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=142)
Your max_length is set to 500, but your input_length is only 472. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=236)
Your max_length is set to 500, but your input_length is only 404. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=202)
Your max_length is set to 500, but your input_length is only 438. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=219)
Your max_length is set to 500, but your input_length is only 325. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=162)
Your max_length is set to 500, but your input_length is only 404. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=202)
Your max_length is set to 500, but your input_length is only 368. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=184)
Your max_length is set to 500, but your input_length is only 322. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=161)
Your max_length is set to 500, but your input_length is only 360. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=180)
Your max_length is set to 500, but your input_length is only 353. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=176)
Your max_length is set to 500, but your input_length is only 364. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=182)
Your max_length is set to 500, but your input_length is only 451. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=225)
Your max_length is set to 500, but your input_length is only 256. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=128)
Your max_length is set to 500, but your input_length is only 412. Since this is a su
```

```
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=206)
Your max_length is set to 500, but your input_length is only 295. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=147)
Your max_length is set to 500, but your input_length is only 358. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=179)
Your max_length is set to 500, but your input_length is only 321. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=160)
Your max_length is set to 500, but your input_length is only 353. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=176)
Your max_length is set to 500, but your input_length is only 376. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=188)
Your max_length is set to 500, but your input_length is only 388. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=194)
Your max_length is set to 500, but your input_length is only 339. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=169)
Your max_length is set to 500, but your input_length is only 327. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=163)
Your max_length is set to 500, but your input_length is only 336. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=168)
Your max_length is set to 500, but your input_length is only 439. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=219)
Your max_length is set to 500, but your input_length is only 354. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=177)
Your max_length is set to 500, but your input_length is only 399. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=199)
Your max_length is set to 500, but your input_length is only 342. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=171)
Your max_length is set to 500, but your input_length is only 280. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=140)
Your max_length is set to 500, but your input_length is only 305. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=152)
Your max_length is set to 500, but your input_length is only 348. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=174)
Your max_length is set to 500, but your input_length is only 483. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=241)
Your max_length is set to 500, but your input_length is only 437. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=218)
```

```
Your max_length is set to 500, but your input_length is only 358. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=179)
Your max_length is set to 500, but your input_length is only 322. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=161)
Your max_length is set to 500, but your input_length is only 410. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=205)
Your max_length is set to 500, but your input_length is only 357. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=178)
Your max_length is set to 500, but your input_length is only 484. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=242)
Your max_length is set to 500, but your input_length is only 314. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=157)
Your max_length is set to 500, but your input_length is only 392. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=196)
Your max_length is set to 500, but your input_length is only 298. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=149)
Your max_length is set to 500, but your input_length is only 343. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=171)
Your max_length is set to 500, but your input_length is only 283. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=141)
Your max_length is set to 500, but your input_length is only 403. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=201)
Your max_length is set to 500, but your input_length is only 297. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=148)
Your max_length is set to 500, but your input_length is only 403. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=201)
Your max_length is set to 500, but your input_length is only 357. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=178)
Your max_length is set to 500, but your input_length is only 426. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=213)
Your max_length is set to 500, but your input_length is only 360. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=180)
Your max_length is set to 500, but your input_length is only 410. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=205)
Your max_length is set to 500, but your input_length is only 396. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=198)
Your max_length is set to 500, but your input_length is only 313. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
```

```
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=156)
Your max_length is set to 500, but your input_length is only 314. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=157)
Your max_length is set to 500, but your input_length is only 459. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=229)
Your max_length is set to 500, but your input_length is only 396. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=198)
Your max_length is set to 500, but your input_length is only 376. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=188)
Your max_length is set to 500, but your input_length is only 315. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=157)
Your max_length is set to 500, but your input_length is only 293. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=146)
Your max_length is set to 500, but your input_length is only 477. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=238)
Your max_length is set to 500, but your input_length is only 343. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=171)
Your max_length is set to 500, but your input_length is only 372. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=186)
Your max_length is set to 500, but your input_length is only 317. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=158)
Your max_length is set to 500, but your input_length is only 288. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=144)
Your max_length is set to 500, but your input_length is only 356. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=178)
Your max_length is set to 500, but your input_length is only 317. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=158)
Your max_length is set to 500, but your input_length is only 223. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=111)
Your max_length is set to 500, but your input_length is only 287. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=143)
Your max_length is set to 500, but your input_length is only 176. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=88)
Your max_length is set to 500, but your input_length is only 432. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=216)
Your max_length is set to 500, but your input_length is only 455. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=227)
Your max_length is set to 500, but your input_length is only 256. Since this is a su
```

```
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=128)
Your max_length is set to 500, but your input_length is only 418. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=209)
Your max_length is set to 500, but your input_length is only 317. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=158)
Your max_length is set to 500, but your input_length is only 246. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=123)
Your max_length is set to 500, but your input_length is only 278. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=139)
Your max_length is set to 500, but your input_length is only 397. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=198)
Your max_length is set to 500, but your input_length is only 322. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=161)
Your max_length is set to 500, but your input_length is only 332. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=166)
Your max_length is set to 500, but your input_length is only 289. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=144)
Your max_length is set to 500, but your input_length is only 243. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=121)
Your max_length is set to 500, but your input_length is only 253. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=126)
Your max_length is set to 500, but your input_length is only 434. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=217)
Your max_length is set to 500, but your input_length is only 342. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=171)
Your max_length is set to 500, but your input_length is only 362. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=181)
Your max_length is set to 500, but your input_length is only 348. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=174)
Your max_length is set to 500, but your input_length is only 491. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=245)
Your max_length is set to 500, but your input_length is only 370. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=185)
Your max_length is set to 500, but your input_length is only 263. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=131)
Your max_length is set to 500, but your input_length is only 409. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=204)
```

```
Your max_length is set to 500, but your input_length is only 334. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=167)
Your max_length is set to 500, but your input_length is only 381. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=190)
Your max_length is set to 500, but your input_length is only 327. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=163)
Your max_length is set to 500, but your input_length is only 283. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=141)
Your max_length is set to 500, but your input_length is only 418. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=209)
Your max_length is set to 500, but your input_length is only 269. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=134)
Your max_length is set to 500, but your input_length is only 290. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=145)
Your max_length is set to 500, but your input_length is only 306. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=153)
Your max_length is set to 500, but your input_length is only 354. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=177)
Your max_length is set to 500, but your input_length is only 379. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=189)
Your max_length is set to 500, but your input_length is only 347. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=173)
Your max_length is set to 500, but your input_length is only 411. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=205)
C:\Users\Minh UBC\miniconda3\Lib\site-packages\transformers\pipelines\base.py:1157:
UserWarning: You seem to be using the pipelines sequentially on GPU. In order to max
imize efficiency please use a dataset
  warnings.warn(
Your max_length is set to 500, but your input_length is only 295. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=147)
Writing csv output...
Output csv file completed
Output txt file completed
Loading file =  PLOS_test.jsonl
n rows = 142
Extracting abstract...
Making summaries....
```

```
Your max_length is set to 500, but your input_length is only 477. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=238)
Your max_length is set to 500, but your input_length is only 484. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=242)
Your max_length is set to 500, but your input_length is only 488. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=244)
Your max_length is set to 500, but your input_length is only 464. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=232)
Your max_length is set to 500, but your input_length is only 344. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=172)
Your max_length is set to 500, but your input_length is only 446. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=223)
Your max_length is set to 500, but your input_length is only 297. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=148)
Your max_length is set to 500, but your input_length is only 475. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=237)
Your max_length is set to 500, but your input_length is only 473. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=236)
Your max_length is set to 500, but your input_length is only 460. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=230)
Your max_length is set to 500, but your input_length is only 272. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=136)
Your max_length is set to 500, but your input_length is only 440. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=220)
Your max_length is set to 500, but your input_length is only 301. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=150)
Your max_length is set to 500, but your input_length is only 271. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=135)
Your max_length is set to 500, but your input_length is only 496. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=248)
Your max_length is set to 500, but your input_length is only 471. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=235)
Your max_length is set to 500, but your input_length is only 280. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=140)
Your max_length is set to 500, but your input_length is only 440. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=220)
Your max_length is set to 500, but your input_length is only 387. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
```

```
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=193)
Your max_length is set to 500, but your input_length is only 482. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=241)
Your max_length is set to 500, but your input_length is only 468. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=234)
Your max_length is set to 500, but your input_length is only 415. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=207)
Your max_length is set to 500, but your input_length is only 438. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=219)
Your max_length is set to 500, but your input_length is only 371. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=185)
Your max_length is set to 500, but your input_length is only 284. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=142)
Your max_length is set to 500, but your input_length is only 473. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=236)
Your max_length is set to 500, but your input_length is only 450. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=225)
Your max_length is set to 500, but your input_length is only 493. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=246)
Your max_length is set to 500, but your input_length is only 486. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=243)
Your max_length is set to 500, but your input_length is only 407. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=203)
Your max_length is set to 500, but your input_length is only 212. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=106)
Your max_length is set to 500, but your input_length is only 395. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=197)
Your max_length is set to 500, but your input_length is only 340. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=170)
Your max_length is set to 500, but your input_length is only 313. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=156)
Your max_length is set to 500, but your input_length is only 471. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=235)
Your max_length is set to 500, but your input_length is only 480. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=240)
Your max_length is set to 500, but your input_length is only 437. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=218)
Your max_length is set to 500, but your input_length is only 459. Since this is a su
```

```
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=229)
Your max_length is set to 500, but your input_length is only 292. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=146)
Your max_length is set to 500, but your input_length is only 489. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=244)
Your max_length is set to 500, but your input_length is only 260. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=130)
Your max_length is set to 500, but your input_length is only 340. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=170)
Your max_length is set to 500, but your input_length is only 216. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=108)
Your max_length is set to 500, but your input_length is only 465. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=232)
Your max_length is set to 500, but your input_length is only 446. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=223)
Your max_length is set to 500, but your input_length is only 370. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=185)
Your max_length is set to 500, but your input_length is only 381. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=190)
Your max_length is set to 500, but your input_length is only 255. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=127)
Your max_length is set to 500, but your input_length is only 189. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=94)
Your max_length is set to 500, but your input_length is only 291. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=145)
Your max_length is set to 500, but your input_length is only 177. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=88)
Your max_length is set to 500, but your input_length is only 257. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=128)
Your max_length is set to 500, but your input_length is only 346. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=173)
Your max_length is set to 500, but your input_length is only 472. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=236)
Your max_length is set to 500, but your input_length is only 312. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=156)
Your max_length is set to 500, but your input_length is only 471. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=235)
```

```
Your max_length is set to 500, but your input_length is only 429. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=214)
Your max_length is set to 500, but your input_length is only 482. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=241)
Your max_length is set to 500, but your input_length is only 366. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=183)
Your max_length is set to 500, but your input_length is only 404. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=202)
Your max_length is set to 500, but your input_length is only 266. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=133)
Your max_length is set to 500, but your input_length is only 404. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=202)
Your max_length is set to 500, but your input_length is only 371. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=185)
Your max_length is set to 500, but your input_length is only 327. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=163)
Your max_length is set to 500, but your input_length is only 277. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=138)
Your max_length is set to 500, but your input_length is only 459. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=229)
Your max_length is set to 500, but your input_length is only 252. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=126)
Your max_length is set to 500, but your input_length is only 332. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=166)
Your max_length is set to 500, but your input_length is only 352. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=176)
Your max_length is set to 500, but your input_length is only 413. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=206)
Your max_length is set to 500, but your input_length is only 417. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=208)
Your max_length is set to 500, but your input_length is only 286. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=143)
Your max_length is set to 500, but your input_length is only 383. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=191)
Your max_length is set to 500, but your input_length is only 385. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=192)
Your max_length is set to 500, but your input_length is only 427. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
```

```
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=213)
Your max_length is set to 500, but your input_length is only 309. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=154)
Your max_length is set to 500, but your input_length is only 264. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=132)
Your max_length is set to 500, but your input_length is only 333. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=166)
Your max_length is set to 500, but your input_length is only 198. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=99)
Your max_length is set to 500, but your input_length is only 218. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=109)
Your max_length is set to 500, but your input_length is only 486. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=243)
Your max_length is set to 500, but your input_length is only 474. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=237)
Your max_length is set to 500, but your input_length is only 498. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=249)
Your max_length is set to 500, but your input_length is only 436. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=218)
Your max_length is set to 500, but your input_length is only 458. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=229)
Your max_length is set to 500, but your input_length is only 225. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=112)
Your max_length is set to 500, but your input_length is only 312. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=156)
Your max_length is set to 500, but your input_length is only 498. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=249)
Your max_length is set to 500, but your input_length is only 281. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=140)
Your max_length is set to 500, but your input_length is only 391. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=195)
Your max_length is set to 500, but your input_length is only 457. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=228)
Your max_length is set to 500, but your input_length is only 301. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=150)
Your max_length is set to 500, but your input_length is only 352. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=176)
Your max_length is set to 500, but your input_length is only 351. Since this is a su
```

```
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=175)
Your max_length is set to 500, but your input_length is only 218. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=109)
Your max_length is set to 500, but your input_length is only 331. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=165)
Your max_length is set to 500, but your input_length is only 274. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=137)
Your max_length is set to 500, but your input_length is only 475. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=237)
Your max_length is set to 500, but your input_length is only 352. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=176)
Your max_length is set to 500, but your input_length is only 432. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=216)
Your max_length is set to 500, but your input_length is only 354. Since this is a su
mmarization task, where outputs shorter than the input are typically wanted, you mig
ht consider decreasing max_length manually, e.g. summarizer('...', max_length=177)
Writing csv output...
Output csv file completed
Output txt file completed
---- All completed----
```