

Chapter_06_요약_summarization.ipynb

transformer · 2024. 8. 7. 16:09

출처

<https://product.kyobobook.co.kr/detail/S000200330771>



트랜스포머를
활용한
자연어 처리

Natural Language
Processing
with Transformers

허강매이스 📖 개발팀이
알려주는 자연어
애플리케이션 구축

NLP 분야 선택
근로소득과 세무
관련 도서

이 책은
자연어 처리 분야
초보자를 위한

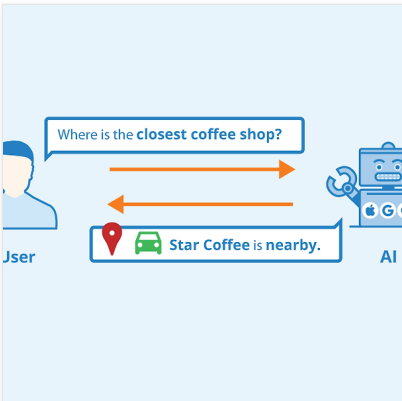
트랜스포머를 활용한 자연어 처리 | ...

트랜스포머를 활용한 자연어 처리 | 챗GPT처럼 자연어를 찢어
같이 알아듣는 트랜스포머 완벽 해부트랜스포머는 우리 주변...

product.kyobobook.co.kr

아래그림 설명

출처:



User: Where is the closest coffee shop?

AI: Star Coffee is nearby.

[NLP] N-gram 언어 모델의 개념, 종...

목차 1. N-gram 개념 2. N-gram 등장 배경 3. N-gram
종류 4. N-gram 기반 단어 예측 5. N-gram의 한계점 1. N...

heytech.tistory.com

3. N-gram 종류


N-gram은 문장을 몇 개의 단어 개수에 따라 나눌지에 따라 종류가 결정됩니다. 아래와 같은 예시 문장을 활용하여 N-gram 언어 모델을 구현했을 때 어떤 결과가 나오는지 알아봄으로써 N-gram 종류에 대해 이해해 봅니다. 문장 부호는 전처리한다고 가정하겠습니다.

"오늘 점심 추천 메뉴: 파스타, 피자"

	모델명	구현 결과
1	Unigram(N=1)	오늘, 점심, 추천, 메뉴, 파스타, 피자
2	Bigram(N=2)	오늘 점심, 점심 추천, 추천 메뉴, 메뉴 파스타, 파스타 피자
3	Trigram(N=3)	오늘 점심 추천, 점심 추천 메뉴, 추천 메뉴 파스타, 메뉴 파스타 피자
4	4-gram(N=4)	오늘 점심 추천 메뉴, 점심 추천 메뉴 파스타, 추천 메뉴 파스타 피자

BLEU

출처: <https://wikidocs.net/31695>



14-03 BLEU Score(Bilingual Eval...

앞서 언어 모델(Language Model)의 성능 측정을 위한 평가 방법으로 퍼플렉시티(perplexity, PPL)를 소개한 바 있습니다...

wikidocs.net

BLEU

모델의 out에서 n-gram으로 따져봤을때 정답에 얼마나 많이 포함되었느냐를 precision의 개념
모델이 생성한 candidate에서의 ngram들이 reference에서 얼마나 많이 포함되었느냐를
precision의 개념
수치적으로 정확도계산
번역이나 요약을 했을때 평가할수 있는 지표

$$y=e^x$$

$$e^{(1-(r/c))} = e^{(1-6/1)} = e^{(1-6)} = e^{(-5)} = 1/100$$

$$e = 2.7$$

$$e^{-5} = \frac{1}{e^5}$$

$$2^3 \quad 2^2 \quad 2^1 \quad 2^0 \quad 2^{-1} \quad 2^{-2}$$

아래그림)

candidate = prediction

reference = 정답

모델이 내놓은 prediction 의 너무 짧으면(the the the the 가아니고 the하나만 있으면 너무짧음) 안되니까 BP를 곱함

penalty를 곱함

출처:<https://wikidocs.net/31695>

```
import pandas as pd
import numpy as np

bleu_metric.add(
    prediction="the the the the the the", reference=["the cat is on
the mat"])
results = bleu_metric.compute(smooth_method="floor", smooth_value=0)
results["precisions"] = [np.round(p, 2) for p in
results["precisions"]]
pd.DataFrame.from_dict(results, orient="index", columns=["Value"])
```

다시 Ref보다 Ca의 길이가 짧은 경우에 패널티를 주는 브레버티 패널티의 이야기로 돌아보겠습니다. 브레버티 패널티는 앞서 배운 BLEU의 식에 곱하는 방식으로 사용합니다. 브레버티 패널티를 줄여서 *BP*라고 하였을 때, 최종 BLEU의 식은 아래와 같습니다.

$$BLEU = BP \times \exp\left(\sum_{n=1}^N w_n \log p_n\right)$$

위의 수식은 패널티를 줄 필요가 없는 경우에는 *BP*의 값이 1이어야 함을 의미합니다. 이를 반영한 *BP*의 수식은 아래와 같습니다.

$$BP = \begin{cases} 1 & \text{if } c > r \\ e^{(1-r/c)} & \text{if } c \leq r \end{cases}$$

c : Candidate의 길이

r : Candidate와 가장 길이 차이가 작은 Reference의 길이

Ref가 1개라면 Ca와 Ref의 두 문장의 길이만을 가지고 계산하면 되겠지만 여기서는 Ref가 여러 개일 때를

아래그림

bigram일때 candidate의 on mat이 reference에 없으니까

bigram일때

prediction : the cat, cat is, is on , on mat (연속된 두개) ->4개 total

이중 reference에 들어간것은 the cat, cat is, is on -> 3개 counts

precision: $3/4 = 0.75$

```

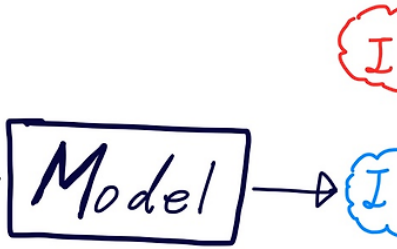
bleu_metric.add(
    prediction="the cat is on mat", reference=["the cat is on the mat"])
results = bleu_metric.compute(smooth_method="floor", smooth_value=0)
results["precisions"] = [np.round(p, 2) for p in results["precisions"]]
pd.DataFrame.from_dict(results, orient="index", columns=["Value"])

```

	Value
score	57.893007
counts	[5, 3, 2, 1]
totals	[5, 4, 3, 2]
precisions	[100.0, 75.0, 66.67, 50.0]
bp	0.818731
sys_len	5
ref_len	6

아래그림

BLEU, ROUGE차이



BLEU Score의 소개와 계산하는 방...

Last Updated on July 23, 2021 Bilingual Evaluation Understudy (BLEU) 딥러닝의 발달로 텍스트를 생성하는...

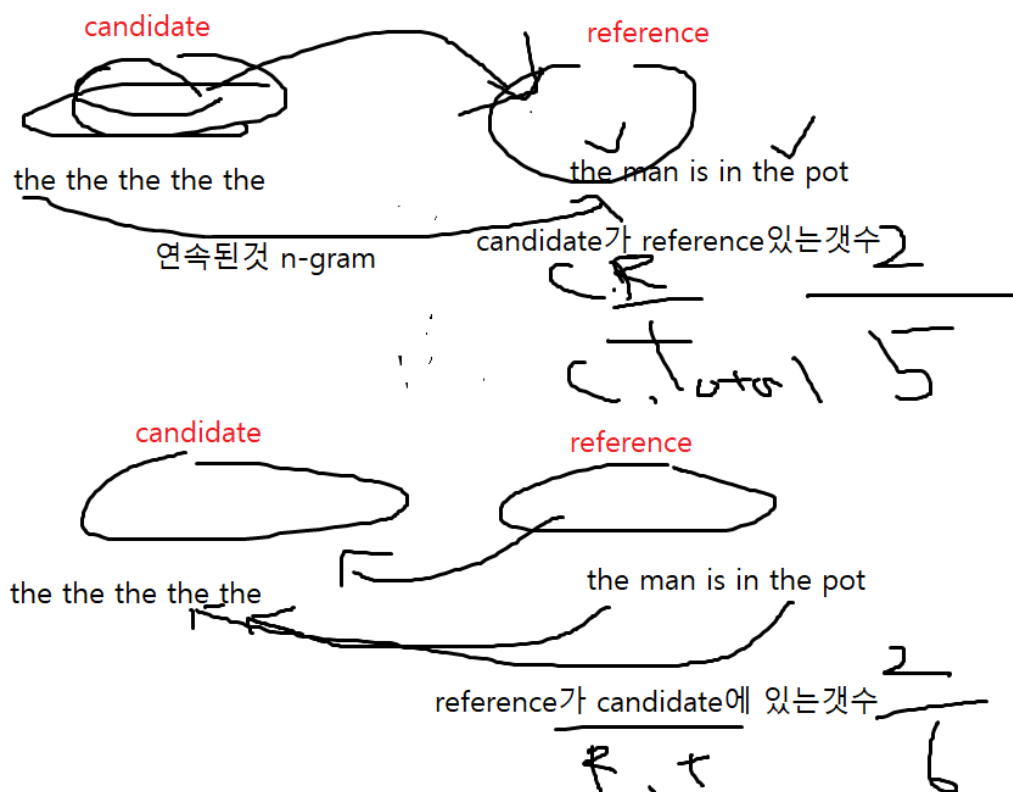
jrc-park.tistory.com

BLEU - 번역작업에 많이 쓰고

- candidate가 reference에 있는갯수

ROUGE - 요약에 많이쓰임

- reference가 candidate에 있는갯수



아래그림)

gpt2는 요약작업에서 pretrain만 되어있어서 finetuning된 bart, pegasus 가 훨씬 요약 잘해서 rouge score 가 높다.

```

+ 코드 + 텍스트
records = []
rouge_names = ["rouge1", "rouge2", "rougeL", "rougeLsum"]

for model_name in summaries:
    rouge_metric.add(prediction=summaries[model_name], reference=reference)
    score = rouge_metric.compute()
    # rouge_dict = dict((rn, score[rn].mid.fmeasure) for rn in rouge_names)
    rouge_dict = dict(score)
    records.append(rouge_dict)
pd.DataFrame.from_records(records, index=summaries.keys())

```

	rouge1	rouge2	rougeL	rougeLsum
baseline	0.303571	0.090909	0.214286	0.232143
gpt2	0.212121	0.000000	0.121212	0.212121
t5	0.486486	0.222222	0.378378	0.486486
bart	0.582278	0.207792	0.455696	0.506329
pegasus	0.866667	0.655172	0.800000	0.833333

▼ CNN/DailyMail 데이터셋에서 PEGASUS 평가하기

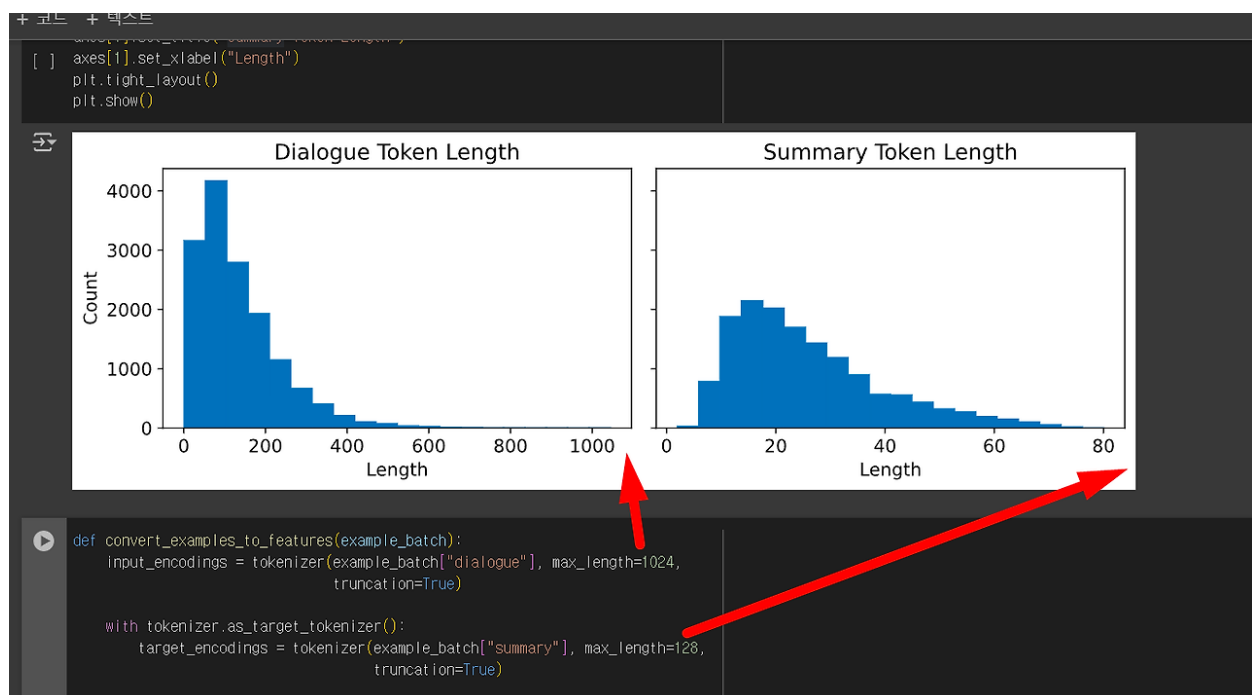
```

[ ] # 이 셀은 노트북 중간부터 실행하기 위한 것입니다.
import matplotlib.pyplot as plt
import pandas as pd
from datasets import load_dataset, load_metric

```

아래그림)

batch size로 제일 긴 max length를 정해서 짧은건 패딩붙이고 기준보다 긴거는 자른다.

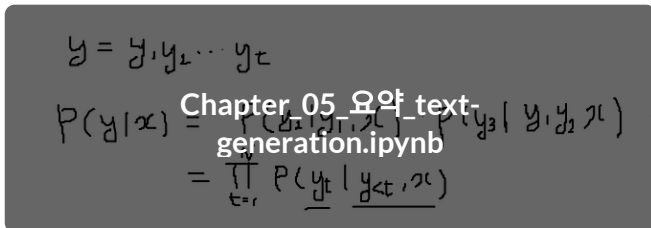
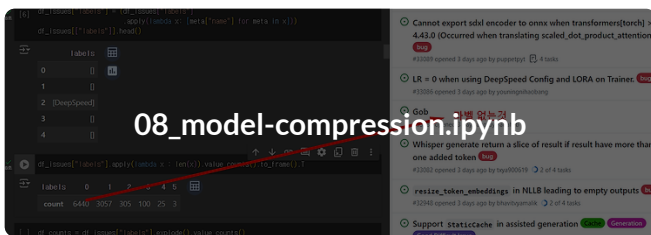


'transformer' 카테고리의 다른 글

08_model-compression.ipynb (0)	2024.08.26
07_question-answering_v2.ipynb (0)	2024.08.16
Chapter_05 요약_text-generation.ipynb (0)	2024.07.31
Chapter_04 요약_multilingual-ner.ipynb (0)	2024.07.26
Chapter_03 요약_transformer 파헤치기 https://nlpinkorean.github.io/illustrated-transform... (0)	2024.07.12

관련글

관련글 더보기



Chapter_04 요약_multilingual-ner.ipynb

RoBERTa						
with BOOKS + WIKI	16GB	8K	100K	93.6/87.3	89.0	95.3
+ additional data (§3.2)	160GB	8K	100K	94.0/87.7	89.3	95.6
+ pretrain longer	160GB	8K	300K	94.4/88.7	90.0	96.1
				90.2	96.4	
BERT _{LARGE}						
with BOOKS + WIKI			1M	90.9/81.8	86.6	93.7
XLNet _{LARGE}						
with BOOKS + WIKI	13GB	256	1M	94.0/87.8	88.4	94.4
+ additional data	126GB	2K	500K	94.5/88.8	89.8	95.6

자연어(NLP)

네이쳐2024 님의 블로그입니다.

구독하기 +

댓글 0



이름

비밀번호

내용을 입력하세요.



등록