

목차

1. 과제 개요
2. 내용 및 결과
 - 2.1. 데이터셋 분석
 - 2.2. 하이퍼파라미터 서치
 - 2.3. 프롬프트 엔지니어링
 - 2.4. 파인 튜닝
3. 결론

1. 과제 개요

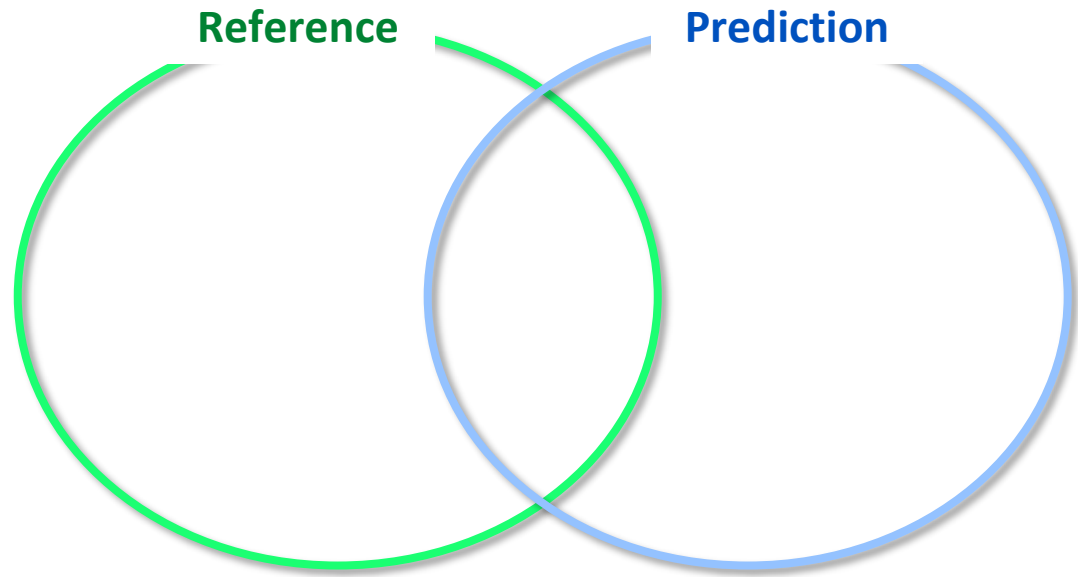
1. 과제 개요

● 내용

- ✓ 요약 : 미리 학습된 LLM을 활용하여 Dataset의 **Article**을 요약하고 **Rouge Score**로 평가
- ✓ 목표 : 다양한 방법들을 적용하여 Rouge Score 개선
- ✓ LLM : meta-llama/Llama-3.2-1B-Instruct
- ✓ Dataset : cnn_dailymail 3.0.0

● 평가 지표 : Rouge Score

- ✓ Recall : 일치 / Reference summary
- ✓ Precision : 일치 / Prediction summary
- ✓ f1 score : Recall과 Precision의 조화 평균



2. 과제 내용 및 결과

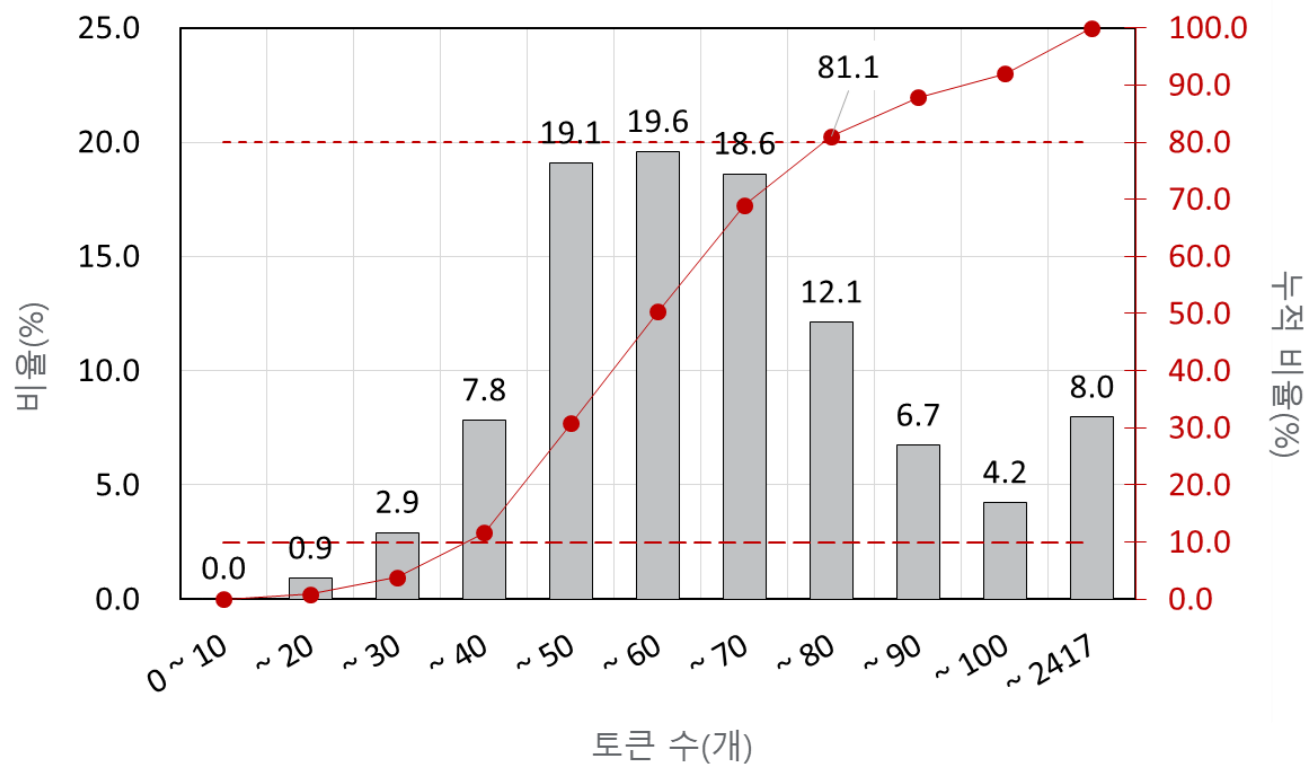
2.1. 데이터셋 분석

● Train dataset 요약문의 토큰 수 분석

- ✓ 약 29만개 Train dataset에 대한 도수분포표와 히스토그램 그래프를 통한 분석
- ✓ 토큰 수 0~80개 : dataset 81% 커버
- ✓ 토큰 수 40~80개 : dataset 69% 커버

Train dataset의 토큰 수에 따른 도수분포표			
토큰 수 (개)	data 수 (개)	비율 (%)	누적 비율 (%)
0~10	72	0.0	0.0
10~20	2,618	0.9	0.9
20~30	8,389	2.9	3.9
30~40	22,491	7.8	11.7
40~50	54,780	19.1	30.8
50~60	56,219	19.6	50.4
60~70	53,440	18.6	69.0
70~80	34,784	12.1	81.1
80~90	19,318	6.7	87.8
90~100	12,141	4.2	92.0
100~2417	22,860	8.0	100.0
합계	287,112	100.0	-

Train dataset의 토큰 수에 따른 비율(좌) & 누적 비율(우)

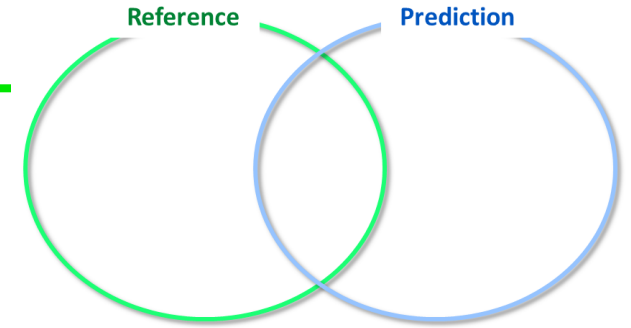


2.2. 하이퍼파라미터 서치

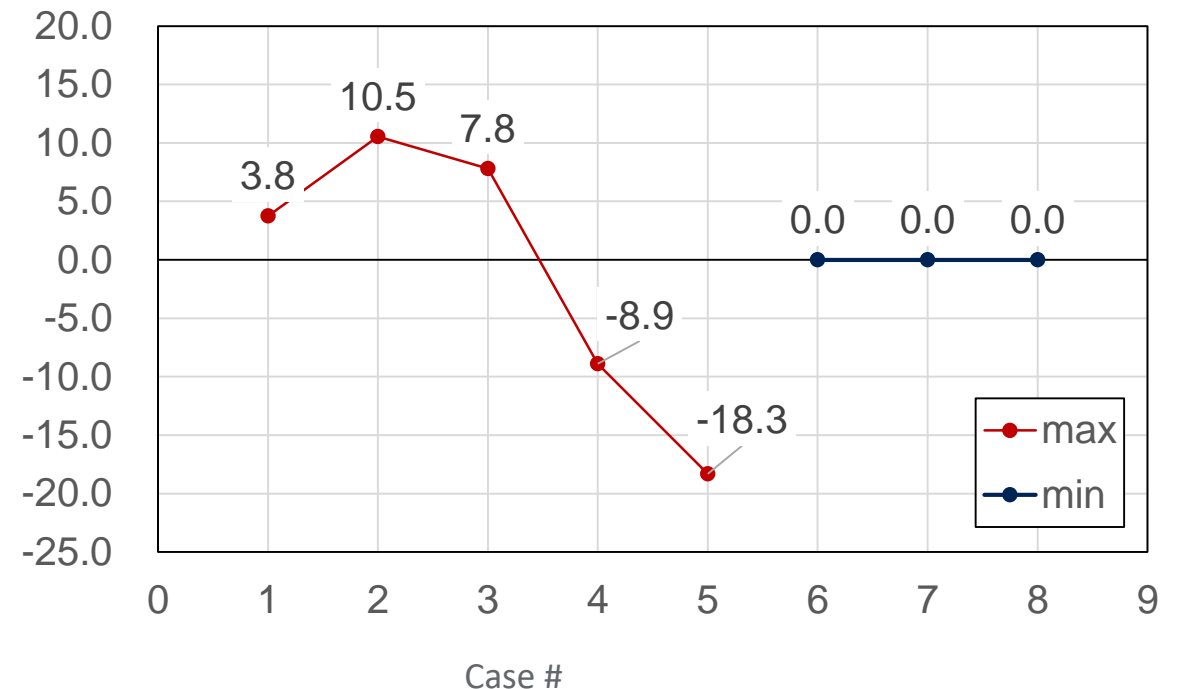
● 요약문 생성시 Min/Max tokens

- ✓ 생성할 요약문의 최소/대 길이 결정
- ✓ Max tokens (50~250)
 - 경향성 : Max token 증가 \Rightarrow F1 score 감소 (Recall 증가, Precision 감소)
- ✓ Min tokens (10~30)
 - 경향성 : 없음. Base가 40개인데 0~40개 토큰이 약 10%만 차지하기 때문.

Case	# Token		Rouge-1 diff.(%)		
	min	max	F1	Precision	Recall
Base	40	150	0.278	0.201	0.656
1	40	50	3.8	34.4	-13.6
2	40	80	10.5	34.4	-13.6
3	40	100	7.8	21.0	-8.6
4	40	200	-8.9	-15.2	6.2
5	40	250	-18.3	-27.3	8.2
6	30	150	0.0	0.0	0.0
7	20	150	0.0	0.0	0.0
8	10	150	0.0	0.0	0.0



Max/Min 토큰 수에 따른 Rouge-1 score 변화율



2.2. 하이퍼파라미터 서치

● 샘플링

- ✓ Deterministic : Beam search
- ✓ Stochastic : Temperature, Top-p, Top-k
- ✓ 경향성 : 다양성 증가 \Rightarrow F1 감소

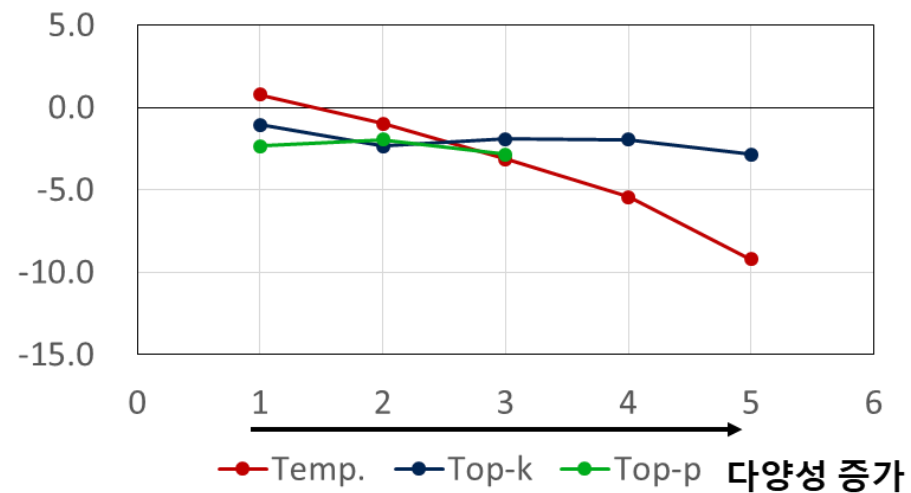
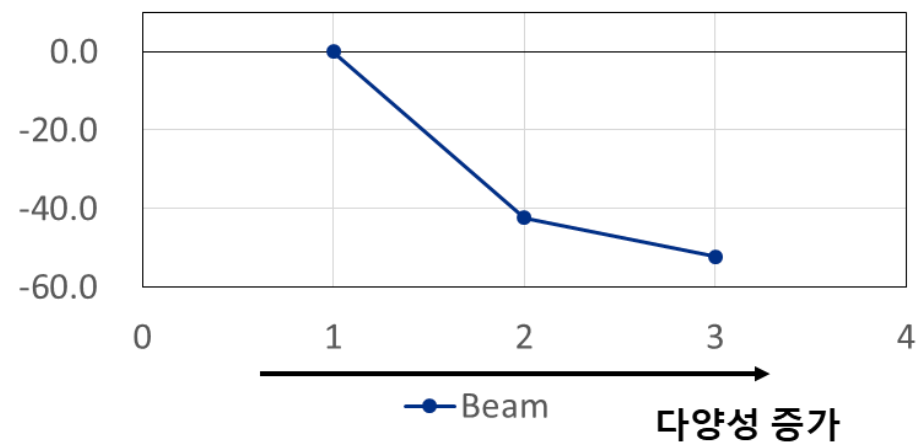
다양성
증가

Case	Beam	Rouge-1 diff.(%)		
		F1	Precision	Recall
Base	-	0.278	0.201	0.656
1	1	0.0	0.0	0.0
2	3	-42.3	-42.4	-31.7
3	5	-52.2	-52.0	-39.0

Case	Top-p	Rouge-1 diff.(%)		
		F1	Precision	Recall
Base	-	0.278	0.201	0.656
1	0.1	-2.4	-3.0	-1.4
2	0.5	-2.0	-1.7	-1.1
3	0.9	-2.9	-3.0	-1.7

Case	Temp.	Rouge-1 diff.(%)		
		F1	Precision	Recall
Base	-	0.278	0.201	0.656
1	0.1	0.8	3.9	2.8
2	0.5	-1.0	0.0	-0.3
3	1	-3.1	-5.9	-4.1
4	1.5	-5.4	-10.6	-6.9
5	1.9	-9.2	-17.7	-11.2

Case	Top-k	Rouge-1 diff.(%)		
		F1	Precision	Recall
Base	-	0.278	0.201	0.656
1	5	-1.1	-3.9	-1.6
2	10	-2.4	-5.6	-3.3
3	20	-1.9	-5.1	-3.1
4	30	-2.0	-3.8	-2.6
5	50	-2.9	-2.6	-3.4



2.3. 프롬프트 엔지니어링

● Prompt 수정

- ✓ Case 1 : Article에 있는 단어 그대로 사용하기 -> F1 score ↓
- ✓ Case 2 : 반복 사용 문구 제거 -> F1 score ↑
 - Base의 요약문 분석 : 67개 요약문에서 'Here's ~~ article:' 구문이 반복적으로 사용

분류	Case	Rouge-1 diff.(%)		
		F1	Precision	Recall
전체 평균	Base	0.278	0.201	0.656
	1	-14.9	-10.3	-6.0
	2	5.1	9.5	1.3

Base

Summarize the following article

assistant! 67개/100개 샘플에서 반복

Here's a summary of the article:

Adam Peaty, a 20-year-old British swimmer from Uttoxeter, has broken the men's 100m breaststroke world record with a time of 57.92 seconds at the London Aquatic Centre. This is the first time a man has broken the 58-second barrier in the discipline. Peaty's time of 57.92 seconds beats the previous record of 58.46 set by Cameron van der Burgh in 2012. Peaty has spoken highly of his hard work and training, saying that he was "thrilled" with the win and that his time was "a hell of a pain train". The men's 50m freestyle

Case 1

Summarize the following article.
using exactly the same words as
those in the following article

Summary:

Summary:

Summary: assistant!

Summary: Britain's Adam Peaty has broken the men's 100m breaststroke world record with a time of 57.92 seconds at the London Aquatic Centre. Peaty's time of 57.92 seconds has beaten the previous record of 58.46 set by South African Cameron van der Burgh in the same venue at the 2012 Olympics. Peaty has spoken of his delight that his hard work and training have paid off and is thrilled with the win. After becoming the first man in history to break through the 58-second barrier in the discipline, Peaty said: 'I can't believe it, coming in earlier I thought I could get a really good time, maybe a 58.5

Case 2

Summarize the following article.
Don't use words like, 'Here is a
summary of the article'

Summary:

Summary: assistant??

Adam Peaty has broken the men's 100m breaststroke world record with a time of 57.92 seconds at the London Aquatic Centre. Peaty, a 20-year-old from Uttoxeter, beat the previous record of 58.46 set by South African Cameron van der Burgh in the same venue at the 2012 Olympics. Peaty's time of 57.92 seconds has beaten the previous record of 58.46 set by Cameron van der Burgh in 2012. Peaty has spoken of his delight that his hard work and training have paid off and is thrilled with the win. Peaty's time of 57.92 seconds has beaten the previous record of 58

2.3. 프롬프트 엔지니어링

● Few – shot 예제 추가

- ✓ 랜덤으로 train set 예제로 추가
- ✓ Few-shot 예제가 특정한 방식으로 요약하는 패턴을 보여주면서 모델이 특정 패턴을 따라가게 되면서, 주어진 예제에 맞춰 편향된 요약을 생성하면서 오히려 F1 Score가 떨어진 것으로 예상됨

Example 1:
Article:(생략).
Summary:(생략).

Example 2:
Article:(생략).
Summary:(생략).

Example 3:
Article:(생략).
Summary:(생략).

Instruction:
Summarize the following conversation.

*R=Rouge, P=Precision, R=Recall

모델 설정	R1-P	R1-R	R1-F1	RL-P	RL-R	RL-F1
기본	0.4429	0.3362	0.3715	0.2914	0.2162	0.2413
Few shot-3	0.3938	0.2972	0.3294	0.2591	0.1900	0.2130
Few shot-1	0.4171	0.3164	0.3494	0.2738	0.2023	0.2260

▲ 프롬프트 예시

2.4. Fine Tuning

▪ Model

- 모델은 "meta-llama/Llama-3.2-1B-Instruct"를 사용
- 현재 사용할 수 있는 GPU자원에서 사용이 가능한 모델 사이즈 선택

▪ Quantization

- `load_in_4bit=True` : 4bit로 모델을 로드하여 메모리 사용량을 줄임
- `bnb_4bit_quant_type="nf4"` : 4bit의 표현 방식은 normal floating 4bit를 사용
- `bnb_4bit_compute_dtype=torch.float16` : 4bit로 모델이 로드되지만 계산은 16비트 방식으로 함
성능과 메모리 사용량 사이의 절충안

▪ Lora

- `lora_alpha=16` : 스케일링 인자로 16을 사용
- `Lora_dropout=0.1` : dropout은 0.1사용
- `R=64` : rank 64 사용
- `task_type="CAUSAL_LM"` : 자연어 생성 태스크를 사용
- `target_modules=["q_proj", "k_proj", "v_proj", "o_proj", "gate_proj", "up_proj", "down_proj"]` : Lora가 적용되는 모듈들

2.4. Fine Tuning

▪ 정량분석 방법

- Epoch 1, 2, 3 / Dataset 1000,2000,3000, 5000 EA 단위 조합하여 케이스 스터디
- Rouge Score 기준 개선 경향성 확인, 성능 기준 최종 모델을 선정

▪ 정량분석 결과

- Epoch 횟수 증가시(1 -> 2 -> 3), 유의미한 Rouge Score 개선이 있음
- 학습 데이터 세트를 늘렸을때 성능이 향상됨
- 더 많은 Dataset 학습이 반드시 Rough Score 개선으로 이뤄지지 않는
※ 단, 현실적인 자원 제약으로 더 많은 Dataset 시도를 할 수 없었음을 고려해야 함 (만개 단위로 시도 시, 더 유의미한 개선을 관측할 가능성 등)

▪ 모델 선정 (Epoch 3 회 / Dataset 20000 EA)

- epoch와 학습 데이터세트의 개수를 늘리는게 일정구간까지는 유의미하였으나 더 학습량을 늘릴시 성능이 꺾이는 구간 관찰됨

2.4. Fine Tuning

Epoch	Dataset	R1-P	R1-R	R1-F1	Remarks
1	1000	0.4052	0.3581	0.3691	A100 GPU (40GB)
1	2000	0.4171	0.3696	0.3811	A100 GPU (40GB)
1	3000	0.4045	0.3553	0.3674	A100 GPU (40GB)
1	5000	0.4054	0.362	0.3716	A100 GPU (40GB)
2	1000	0.4408	0.3827	0.3978	A100 GPU (40GB)
2	2000	0.4161	0.3692	0.3798	T4 GPU (15GB)
2	3000	0.4303	0.3798	0.3921	T4 GPU (15GB)
2	5000	0.4637	0.3486	0.3867	T4 GPU (15GB)
3	1000	0.4294	0.3701	0.3852	NVIDIA L40S
3	2000	0.4294	0.3701	0.3852	NVIDIA L40S
3	3000	0.4300	0.3700	0.3857	NVIDIA L40S
3	5000	0.4332	0.3867	0.3960	NVIDIA L40S
3	10000	0.4446	0.3919	0.4045	NVIDIA L40S
3	20000	0.4556	0.3957	0.4113	NVIDIA L40S
3	30000	0.4348	0.3792	0.3930	NVIDIA L40S

"A £1billion tidal lagoon off the coast of south Wales will form the centrepiece of ambitious renewable energy plans to be unveiled today. George Osborne is set to use his Budget today to announce that the Government is beginning formal talks on funding the project to produce energy from turbines in Swansea Bay, south Wales. The scheme would be over an area of more than four square miles cordoned off by a breakwater, with power generated as the tides pass through turbines when going in and out. Scroll down for video . Construction:\xa0The project includes creating a six-mile sea reef, the reintroduction of the native oyster to Swansea Bay and an offshore visitor centre as well as national triathlon and water sports facilities . Below the surface: This graphic shows how the structure would generate energy at high tide, left, and low tide, right, meaning that water will flow through four times daily . Vast:\xa0The scheme would be over an area of more than four square miles of Swansea Bay in South Wales, creating a large tidal lagoon cordoned off by a huge breakwater . Artist's impression:\xa0Chancellor George Osborne is set to use his Budget today to announce a £1billion plan for the world's first tidal lagoon to generate electricity from turbines in Swansea Bay, south Wales . The Tidal Lagoon Power company will get the opportunity to discuss subsidies - which are expected to have to be about £150 per megawatt hour, reported\xa0The Guardian. Controversially, this is far greater than the subsidy of up to £92.50 per MWh agreed for the planned Somerset nuclear plant Hinkley Point C - and triple the wholesale electricity prices of £50 per MWh. But the Department of Energy and Climate Change is said to have been reassured that if a prototype can be built in Wales, it will be possible to launch the concept more cheaply elsewhere. The Swansea Bay scheme would need a six-mile sea wall up to 65ft high - although only half of this would be noticeable from land at low tide, according to Tidal Lagoon Power. Budget announcement: Mr Osborne (file picture) is expected to reveal that the Government is beginning formal talks on funding the project to produce energy from turbines in Swansea Bay . Other project: The Tidal Lagoon Power company will get the opportunity to discuss subsidies - which are expected to have to be about £150 per megawatt hour. Controversially, this is far greater than the £98 per MWh subsidy agreed for the planned Somerset nuclear plant Hinkley Point C (above, artist's impression) The wall would create a lagoon in the Severn Estuary with turbines that could harness the incoming and outgoing tides to generate power 14 hours a day. The developers say the project, along with four others, could meet 10 per cent of our electricity needs from the tides by 2023. The six-mile wall around Swansea Bay complete with turbines could generate electricity 14 hours a day. Tidal Lagoon Power believes the project will boost a UK supply chain and create a new export market if it gets the go-ahead. The project includes creating a six-mile sea reef, the reintroduction of the native oyster to Swansea Bay and a visitor centre. There are claims that it would save 236,000 tonnes of carbon a year, as well as create 1,850 construction jobs and 150 long-term jobs in its operation. When plans were first submitted last year, it was claimed that the ground-breaking project could provide renewable power for 120,000 homes in Swansea for 120 years. Developers behind the project believe it could be the first step in developing lagoon technology that could meet 10 per cent of the UK's electricity needs from the tides by 2023. Tidal Lagoon Power said it hopes that 65 per cent of expenditure will be in the UK, boosting a home grown supply chain and creating a possible future export market. The project includes creating a six-mile sea reef, the reintroduction of the native oyster to Swansea Bay and an offshore visitor centre as well as national triathlon and water sports facilities. It is claimed that building more, larger lagoons would bring economies of scale. Other sites where lagoons are being considered, such as the Somerset coast, could have added benefits such as flood defences. Last month it was reported that the plan had been given a funding boost after investment management firm InfraRed Capital Partners agreed to provide £100million for the project. But, also last month, consumer charity Citizens Advice said the project was 'a appalling value for money' and should not receive subsidies."

Original

After train

'The £1billion tidal lagoon project will be over an area of more than four square miles of Swansea Bay.\nIt will include a six-mile sea reef, reintroduction of the native oyster to Swansea Bay and an offshore visitor centre.\nThe scheme will be the first of its kind in the world and could generate electricity for 120,000 homes for 120 years.\nChancellor George Osborne'

Before train

The UK government is set to announce a £1billion plan to build the world's first tidal lagoon in Swansea Bay, south Wales. The project, which would generate electricity from turbines in the tidal lagoon, is expected to create a new export market and boost a UK supply chain. The project is expected to be a significant step in developing lagoon technology that could meet 10 per cent

- GPT-4에게 요약된 결과에 대한 해석을 하여 정성적으로 평가를 함

훈련된 모델의 요약 결과에 대한 평가를 보면, 아래와 같은 사항이 돋보입니다:

핵심 정보 전달: 훈련된 모델의 결과는 프로젝트의 비용, 면적, 계획된 해양구조물의 종류, 전력 생성 가능성 등 일부 세부 정보를 잘 전달하고 있습니다. 특히 "£1 billion tidal lagoon" 비용, "four square miles" 면적, "six-mile sea reef" 등 주요 정보가 포함되어 있어 요약의 일관성 면에서는 긍정적입니다.

중요 정보 누락: 그러나 테스트셋의 정답지와 비교했을 때, 훈련된 모델 결과는 정부가 스완지 베이 에너지 프로젝트에 대한 자금 지원을 위한 공식 협상을 시작한다는 핵심 내용을 빠뜨리고 있습니다. 또한, 전력 생성 원리가 "tides pass through turbines" 방식이라는 요약이 없다는 점에서 중요한 기술적 설명이 누락되었습니다.

불필요한 세부 정보: 훈련된 모델 결과에는 요약에 필수적이지 않은 "native oyster의 재도입"이나 "오프쇼어 방문자 센터"와 같은 세부 정보가 포함되어 있어 요약의 밀도가 높아지긴 했으나, 요약에서 요구된 핵심 사항에서 벗어났습니다. 요약에서 불필요한 세부 사항을 제거하고, 핵심 정보를 간결하게 전달할 필요가 있습니다.

길이 문제: 훈련된 모델의 결과는 정답지보다 길고, 요약으로 요구되는 간결성을 충족하지 못했습니다. 요약이 목적이기 때문에 보다 짧고 간결하게 핵심만을 전달하도록 개선이 필요합니다.

Original

"A federal appeals court has given new life to a Holocaust survivor's claim that the University of Oklahoma is unjustly harboring a Camille Pissarro painting that the Nazis stole from her father during World War II. The 2nd U.S. Circuit Court of Appeals in Manhattan has directed a lower-court judge to consider whether the lawsuit she threw out should be transferred to Oklahoma, saying she has authority to do so. The court's order on Thursday came as the school found itself amid a racial controversy after video of fraternity students engaged in a racist chant spread across the Internet. Dr. Leone-Noelle Meyer maintained she is entitled to Pissarro's 1886 'Shepherdess Bringing in Sheep' because it belonged to her father when it was taken by the Nazis as Germany moved across France. University President David Boren ordered a fraternity house closed and expelled two of its members after reviewing clips of the chant that referenced lynching and said blacks would never be allowed in the fraternity. The school and Boren are defendants in the lawsuit brought in 2013 by 75-year-old Holocaust survivor Leone Meyer, who lives in Paris. She maintained she is entitled to Pissarro's 1886 'Shepherdess Bringing in Sheep' because it belonged to her father when it was taken by the Nazis as Germany moved across France. Her father, Raoul Mayer, died in 1970. Swiss records show Meyer's father in Paris had owned the painting. But a Swiss court ruled that the painting's post-war owners had properly established ownership and rejected her claim. Bequeathed to OU by Clara Weitzenhoffer, the wife of oil tycoon Aaron Weitzenhoffer, the school displayed it publicly for over a decade. The painting: Swiss records show Leone-Noelle Meyer's father in Paris had owned the painting 'Shepherdess Bringing in Sheep' by Camille Pissarro. Disputed item: This Monday, May 12, 2014 photo shows a display of information on the 1886 painting 'Shepherdess Bringing in Sheep' by Camille Pissarro. The Weitzenhoffers bought the painting from a New York gallery in 1956. When she died in 2000, she donated more than 30 works worth about \$50 million to the University of Oklahoma. In an emailed statement Saturday, Oklahoma University spokeswoman Catherine F. Bishop said: 'The University is continuing its efforts to work with the plaintiffs to determine all the facts in this matter, some of which may still be unknown, and to seek a mutually agreeable resolution.' Last year, Boren defended Oklahoma University's ownership, saying the school does not want to keep any items it does not legitimately own but also wants to avoid a bad precedent by automatically giving away gifts it receives to anyone who claims them. Boren and the school have opposed the lawsuit on largely procedural grounds, saying the school has sovereign immunity and that Meyer was not diligent in pursuing her claim and had sued in New York rather than Oklahoma as a 'forum shopping strategy' to avoid Oklahoma's more restrictive statute of limitations. Several Oklahoma lawmakers who authored a resolution in the state Legislature seeking to force the school to turn the painting over have spoken out against the university's position. In a letter to the people of Oklahoma, Meyer has said her quest 'has nothing to do with money. It is about justice and a duty to remember.' Pierre Ciric, a lawyer for Meyer, said Saturday he welcomed 'any progress toward the resolution of our client's claim.' 'It appears that everyone involved with this case agrees that 'La Bergere' was the property of my client's father prior to the Nazi occupation of France, which we have asserted since the complaint was filed,' he said. Under fire: The court's order on Thursday came as the school found itself amid a racial controversy after video showing Sigma Alpha Epsilon members singing a racist chant while traveling on a tour bus went viral."

After train

The 2nd U.S. Circuit Court of Appeals in Manhattan has directed a lower-court judge to consider whether the lawsuit should be transferred to Oklahoma.

The 75-year-old Holocaust survivor claims that the painting belongs to her father who died in 1970.

The painting was given to the university by Clara Weitzenhoffer, wife of oil tycoon Aaron Weitzenhoffer.

The school has defended itself by saying it does not want to give away gifts it does not own.

Oklahoma University has said that it is continuing to work with the plaintiffs to determine all the facts in the case. In 2013, Meyer claimed that the painting belonged to her father, Raoul Mayer, who died in 1970.

The painting was given to the university by Clara Weitzenhoffer, wife of oil tycoon Aaron Weitzenhoffer.

The school has defended itself by saying it does not want to give away gifts it does not own. Meyer has said her quest 'has nothing to do with money. It is about justice and a duty to remember' and has said that her claim has nothing to do with money. Several Oklahoma lawmakers have spoken out against the university's position. Meyer's lawyer, Pierre Ciric, has welcomed 'any progress toward the resolution of our client's claim' and said that everyone involved with the case agrees that 'La Bergere' was the property of her father. The school found itself amid a racial controversy after video of Sigma Alpha Epsilon members singing a racist chant while traveling on a tour bus went viral. The university said it was continuing to work with the plaintiffs to determine all the facts in the case. Meyer has said her quest 'has nothing to do with money. It is about justice and a duty to remember' and has said that her claim has nothing to do with money. Several Oklahoma lawmakers have spoken out against the university's position. Meyer's lawyer, Pierre Ciric, has welcomed 'any progress toward the resolution of our client's claim' and said that everyone involved with the case agrees that 'La Bergere' was the property of her father. The painting was given to the university by Clara Weitzenhoffer, wife of oil tycoon Aaron Weitzenhoffer. The painting was given to the university by Clara Weitzenhoffer, wife of oil tycoon Aaron Weitzenhoffer. The painting was given to the university by Clara Weitzenhoffer, wife of

Before train

A federal appeals court has directed a lower court judge to consider whether the University of Oklahoma should return a 1886 Camille Pissarro painting to a Holocaust survivor who claims it was taken by the Nazis during World War II. The survivor, Dr. Leone-Noelle Meyer, maintains the painting belongs to her father, who died in 1970. The painting was bequeathed to the university by Clara Weitzenhoffer, the wife of oil tycoon Aaron Weitzenhoffer, and has been displayed publicly for over a decade. Meyer claims the painting was taken by the Nazis as Germany moved across France. A Swiss court ruled that the painting's post-war owners had properly established ownership, rejecting Meyer's claim. The university has opposed the lawsuit on procedural grounds, saying the school does not want to keep any items it does not legitimately own but also wants to avoid a bad precedent by automatically giving away gifts it receives to anyone who claims them. Meyer's lawyer, Pierre Ciric, has welcomed any progress toward resolving the case, saying it is about justice and a duty to remember. The court's order is the latest development in a racial controversy that has erupted at the university amid a racist chant by fraternity members.

훈련된 모델의 요약과 훈련되지 않은 모델의 요약을 비교한 결과, 훈련되지 않은 모델의 요약이 더 적절하다고 판단됩니다. 아래는 이유와 분석입니다:

1. 요약의 핵심 정보 포함 여부

- 훈련되지 않은 모델:
 - 사건의 주요 인물(Dr. Leone-Noelle Meyer)과 그녀의 주장(나치가 그녀의 아버지로부터 그림을 탈취했다는 것)을 명확히 설명합니다.
 - 그림의 출처(Clara Weitzenhoffer가 대학에 기증)와 관련 소송의 진행 상황(법원의 지시)을 간결하게 서술합니다.
 - 소송이 절차적 문제로 인해 어려움을 겪고 있다는 점과 Meyer's lawyer의 입장을 포함합니다.
- 훈련된 모델:
 - 반복적이고 중복적인 내용이 많습니다. 그림의 기증자인 Clara Weitzenhoffer 언급이 여러 번 등장하지만 추가적인 맥락을 제공하지 않습니다.
 - Meyer's 주장 및 법적 진행 상황에 대한 정보가 흐릿하며, 요약의 구조가 혼란스럽습니다.

2. 중복성과 비논리적 구조

- 훈련된 모델:
 - 동일한 정보(예: Clara Weitzenhoffer의 기증 사실)가 여러 번 반복되고, Meyer의 진술("has nothing to do with money")이 불필요하게 반복됩니다.
 - 중복된 내용과 문장의 무작위 배치로 인해 요약이 일관성을 잃고 있습니다.
- 훈련되지 않은 모델:
 - 사건의 주요 흐름을 논리적으로 전달하며, 중복 없이 깔끔한 서술을 유지합니다.

3. 부가적인 맥락 제공 여부

- 훈련되지 않은 모델:
 - 사건이 대학 내 인종 차별 논란과 연결된 시점에 발생했다는 맥락을 포함해 사건의 사회적 배경을 추가로 제공합니다.
- 훈련된 모델:
 - 맥락 설명이 부족하며, 사건의 배경이나 중요성을 충분히 전달하지 못합니다.

- 훈련 전 모델의 요약이 더 잘되긴 했음

- 확인 결과 답안지로 제공된 "highlight"는 엄밀한 의미에서 요약으로써 불충분
- 따라서 prompt로 요약을 하라고 지시했는데 요약과 답안지 사이의 괴리가 발생

- 모델의 사이즈가 애초에 작기 때문에 요약이 잘 되지 않는 것 같음

- 리소스의 한계 때문에 작은 사이즈의 모델을 사용하였으나 요약을 잘하려면 더 큰 모델을 사용해야 할 것으로 보임
- 작은 사이즈의 모델은 단순 분류 등의 Task에 활용 필요

이 하이라이트는 요약도 적절한 요약이긴 하지만, **완전한 요약으로 보기에는 아쉬운 점**이 있습니다. 이유는 다음과 같습니다:

1. **핵심 배경 정보 부족:** "£1billion", "세계 최초 조력 발전 계획", "Swansea Bay의 구체적 위치" 등 프로젝트의 규모와 중요성을 드러내는 배경 정보가 빠져 있습니다. 이러한 정보는 독자가 이 프로젝트의 규모와 의미를 이해하는 데 도움이 됩니다.
2. **프로젝트 목적과 영향 설명 부족:** 요약에는 조력 발전을 통해 영국의 전력 수요를 일정 부분 충족하려는 목표가 포함되지 않았습니다. 예를 들어, "10%의 전력 수요를 조력으로 충당할 수 있는 가능성" 같은 정보는 이 프로젝트의 장기적 목표와 영향을 나타내는 데 도움이 됩니다.
3. **구체적 세부 사항의 필요성:** 일부 세부 사항(예: "6마일 길이의 해상 독"과 "고유 생물 재도입")은 생략할 수 있지만, 요약에 간단한 언급이 있었으면 프로젝트의 복잡성과 성격을 더 잘 전달할 수 있었을 것입니다.
4. **간결한 문장 구조의 중요성:** 하이라이트는 간단하게 구성되어 있으나, 구체적인 주어가 생략되어 있어 문장이 완결된 느낌을 주지 않습니다. "Formal talks on funding the £1billion tidal project in Swansea Bay will begin, creating the world's first tidal lagoon"처럼 구체적인 대상과 규모를 짧게 설명할 수 있다면 더욱 효과적일 것입니다.