



2025-하계 세종창의학기제(집중이수제) 주간학습보고서 (4주차)

창의과제	스마트노트				
이름	임홍철	학습기간	7월 21일 ~ 7월 25일		
학번	20011815	학습주차	4주차	학습시간	12
학과(전공)	데이터사이언스학과	과목명	자기주도창의전공 I	수강학점	3
※ 수강학점에 따른 회차별 학습시간 및 10회차 이상 학습 준수					
금주 학습목표	최종적으로 제출할 앱을 어떤 식으로 디자인 할 것인지 결정 및 구현				
학습내용	우리들의 프로그램을 어떤식으로 만들지 디자인 해보고, 어떤 기능이 들어가야 할지 고민해봤습니다. 아래 첨부된 내용에 해당 내용이 있습니다.				
학습방법	인터넷 서칭(github 코드 탐구)				
학습성과 및 목표달성도	우선 초기 구현을 어떻게 할 것인지 크게 뼈대를 잡게 되었다.				
참고자료 및 문헌	https://flask.palletsprojects.com/en/stable/				
내주 계획	설계된 대로 구현에 들어감				

2025 년 7 월 29 일

지도교수 전창재

(인)

1

대충 우리의 프로그램에서 거쳐야 하는 코드들을 기준으로 %잡으면 됨

1.영상,교안 알맞게 왔는지

2.영상 whisper로 변환

3.교안속 이미지 3d가능판별 및 추출

4.동영상 속 장면변화패 이미지 시간대 추출

5.Ilm에 교안과 대사넣고 값 반환

6.기타 자잘한 것들

스마트 노트 (첫 표지에 알맞은 디자인 필요)

강의 영상

첨부된 동영상 파일 이름 출력

Mp4만 가능 파일 첨부

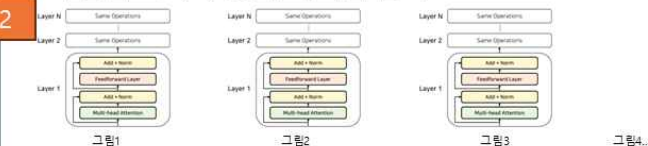
강의 교안

첨부된 교안 파일 이름 출력

Pdf만 가능 파일 첨부

진행도 표시바, 진행 내용과 함께

1 추출된 이미지 3d화 하려면 클릭

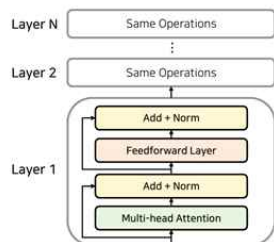


2

해당 페이지 대사 스크립트 전문
(가능하면 여기 대사를 블록 단위로 클릭할 수 있게 하여 클릭하면 아래 타임라인이랑 연동되게)

3 트랜스포머의 동작 원리: 인코더(Encoder)

- 어텐션(Attention)과 정규화(Normalization) 과정을 반복합니다.
- 각 레이어는 서로 다른 파라미터를 가집니다.



Attention is All You Need (NIPS 2017)

나동빈

4

해당 페이지 정보랑 대사스크립트 조합
Gemma 2.5b에 넣어 나온 결과
스크롤 할 수 있게

5

1슬라이드

2슬라이드

3슬라이드

4슬라이드

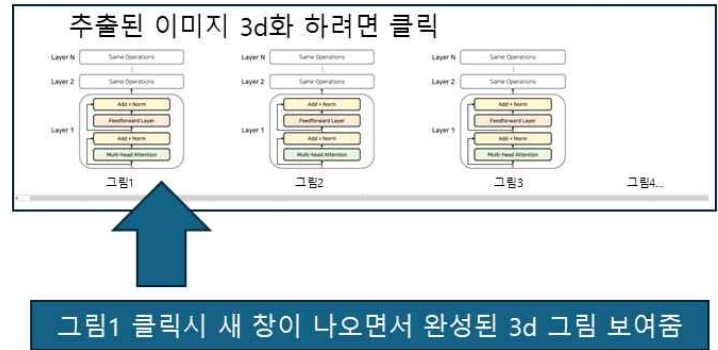
5슬라이드

6슬라이드

7슬라이드

8슬라이드

1 기능 설명



- 그림 그대로 ppt화면에서 해림님의 그림 판별 모델을 이용하여 3d화가 가능한 이미지만 추출하여 가로로 스크롤 가능하게 만들어 클릭하면 팝업형식이나, 새 링크에 만들어진 3d화 모델을 보여준다.

2 기능 설명 – 대사 스크립트 전문

- 텍스트를 넣을때 버튼 형식으로 넣어 클릭하면 5번 기능에서 재생바가 해당 하는 시간으로 변경하여 재생할 수 있게 한다.
- 오른쪽 느낌의 데이터셋을 기준으로 한다.
- 그리고 내용이 길면 이것도 위아래로 스크롤이 가능해야한다.

시간대	대사
0	꼼꼼한 딥러닝 논문 리뷰와 코드 실습. 이번 시간에..
2.92	꼼꼼한 딥러닝 논문 리뷰와 코드 실습. 이번 시간에..
4.6	꼼꼼한 딥러닝 논문 리뷰와 코드 실습. 이번 시간에..
7.32	꼼꼼한 딥러닝 논문 리뷰와 코드 실습. 이번 시간에..
9.22	꼼꼼한 딥러닝 논문 리뷰와 코드 실습. 이번 시간에..
10.42	꼼꼼한 딥러닝 논문 리뷰와 코드 실습. 이번 시간에..

index	start	end	text	punded	sd	page
1	0	2.92	꼼꼼한 딥러	3	0	
2	2.92	4.6	이번 시간	5	0	
3	4.6	7.32	현대 딥러	8	0	
4	7.32	9.22	핵심 야기	10	0	
5	9.22	10.42	트랜스포	11	0	
6	10.42	13	트랜스포	13	0	
7	13	15.3	Attention	16	0	
8	15.3	17.1	논문의 제	18	0	
9	17.1	19.2	트랜스포	20	0	
10	19.2	21.1	이 attenti	22	0	
11	21.1	23.86	가장 예민	24	0	
12	23.86	24.9	알 수 있습	25	0	
13	24.9	26.36	실제로 트	27	0	
14	26.36	27.96	attentionC	28	0	
15	27.96	30.18	전적으로	31	1	
16	30.18	32.28	트랜스포	33	1	
17	32.28	33.42	이해하기	34	1	
18	33.42	35.96	딥러닝 기	36	1	
19	35.96	37.12	확인해 보	38	1	
20	37.12	38.68	2021년 기	39	1	
21	38.68	41.56	최신 자연	42	1	
22	41.56	43.18	이런 트랜	44	1	
23	43.18	44.48	기반으로	45	1	
24	44.48	45.98	최근까지	46	1	
25	45.98	48.06	GPT와 플	49	1	
26	48.06	49.96	이러한 트	50	1	
27	49.96	51.22	작업의 활	52	1	
28	51.22	52.86	좋은 성능	53	1	
29	52.86	54.48	대표적인	55	1	
30	54.48	57.06	트랜스포	58	1	
31	57.06	58.52	보트는 트	59	1	
32	58.52	59.58	인코더 야	60	1	
33	59.58	61.22	이러한 트	61	1	



3 설명

- 그냥 단순히 받은 pdf파일을 이미지로 전환한후에 해당 페이지에 맞게 보여준다.

4 설명 아래와 같이 llm출력값 넣음 (스크롤할 수 있게)

- 트랜스포머의 동작 방식은 입력 값을 여러 개의 인코더 레이어를 거쳐 처리한 후, 인코더의 최종 출력 값을 디코더에 전달하는 방식으로 이루어집니다. 디코더는 이 인코더의 출력 값을 활용하여, 각 출력 단어를 생성할 때 입력 소스 문장의 어떤 단어에 집중해야 하는지를 결정합니다. 디코더 또한 여러 레이어로 구성되며, 각 레이어는 인코더의 최종 출력 값을 입력으로 받아, 단어 정보를 처리하고 상대적인 위치 정보를 인코딩 값을 추가하여 입력합니다. 디코더 레이어는 셀프 어텐션을 통해 문장 내 단어 간의 관계를 학습하고, 인코더-디코더 어텐션을 통해 인코더의 출력 정보를 활용하여 소스 문장의 어떤 단어와 연관성이 있는지 계산합니다. 이러한 어텐션 과정을 통해 인코더에서 나온 출력 결과를 효과적으로 활용하여 번역 결과를 생성하는 것입니다. 특히, 트랜스포머에서 사용되는 어텐션은 여러 개의 헤드를 가진 멀티헤드 어텐션으로, 각 헤드는 동일한 함수로 동작하지만 Query, Key, Value의 사용 방식이 위치에 따라 달라집니다. 예를 들어 디코더의 출력 단어는 Query가 되고, 인코더의 출력 값이 Key와 Value로 사용되어, 각 단어를 출력하기 위해 어떤 정보를 참고해야 하는지를 결정합니다.



5 설명 타임라인

- Mp3를 추출하여 넣어주고 각 슬라이드 별 슬라이싱 하여 그 페이지가 어느 구간인지 보여준다.
- 중간에 있는 주황색공을 움직여서 시간을 조절할 수 있게 한다.
- 왼쪽에는 음소거 버튼과, 시작, 정지 버튼을 만든다.

