



SweetMe

Study with Me

Plow 조

발표자 장다은

구민영, 박찬진, 전우영

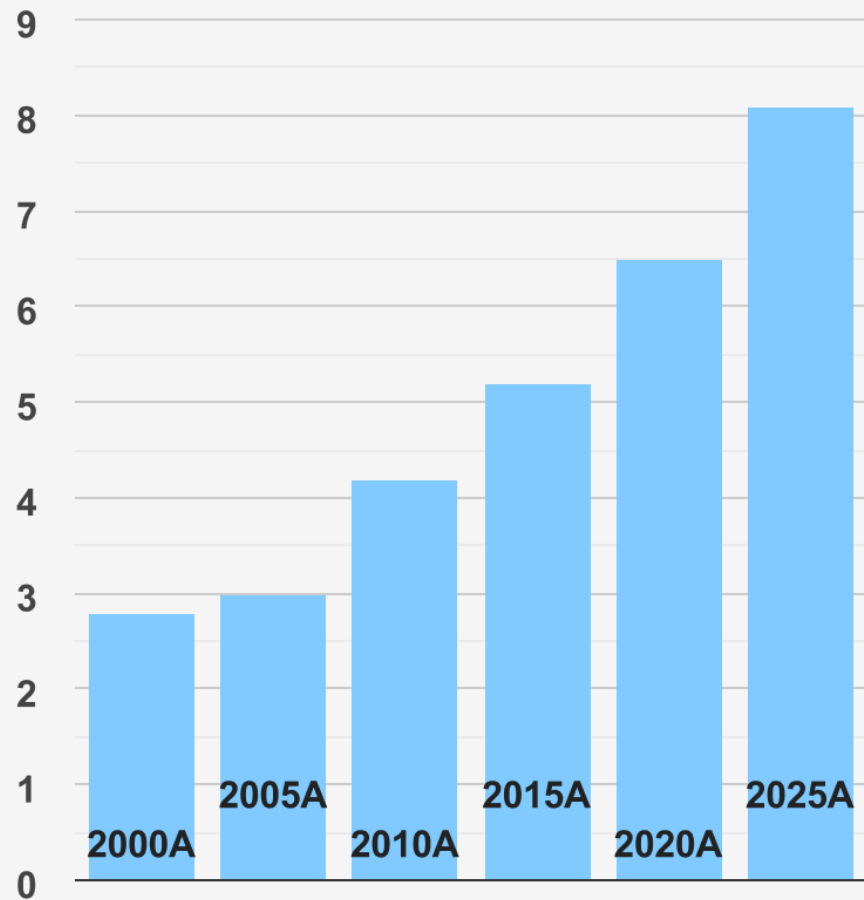
Content

1. 프로젝트 개요
2. 프로젝트 팀 구성 및 역할
3. 프로젝트 진행 프로세스
4. 프로젝트 결과
5. 자체 평가 및 보완

1. 프로젝트 개요 - ① 배경

<세계 교육서비스 시장 규모>

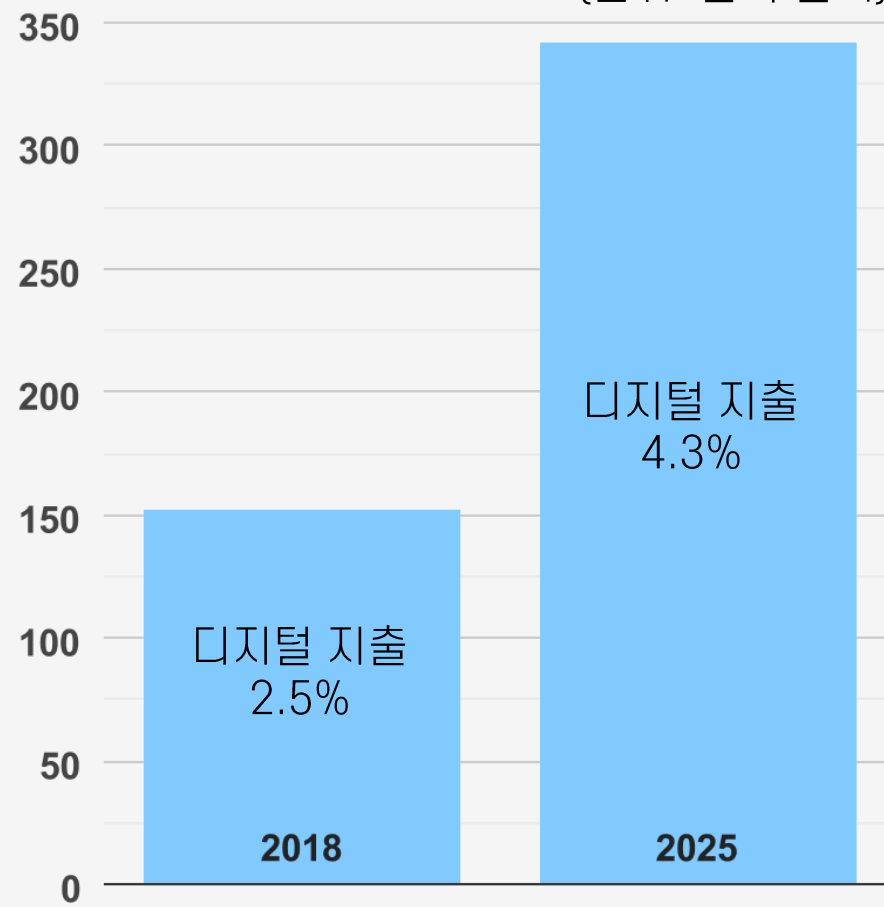
(단위: 조 달러)



출처: HolonIQ (2019.1)

<에듀테크 시장 규모>

(단위: 십억 달러)



1. 프로젝트 개요 - ② 기대 효과



- ✦ 학습자 지원 및 학습 향상
- ✦ 사용자들 간의 협업
- ✦ 확장성과 다양한 타겟 층
- ✦ B2G 사업 참여 가능 (기능 확장 시)

1. 프로젝트 개요 - ① 기술 스택

AI 관련



Hugging face



Google-MLkit

Backend



Firebase



Flask



DataBase



Google cloud

Frontend



Android Studio



Flutter

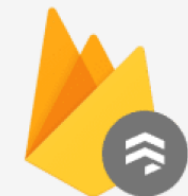
Server



Apache

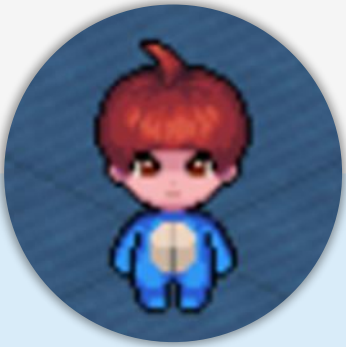


AWS EC2



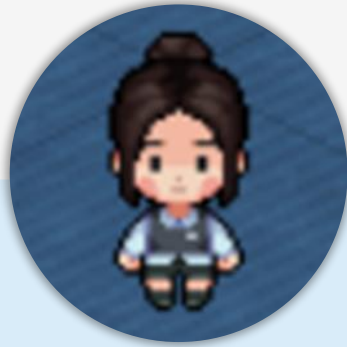
Firestore

2. 프로젝트 팀 구성 및 역할



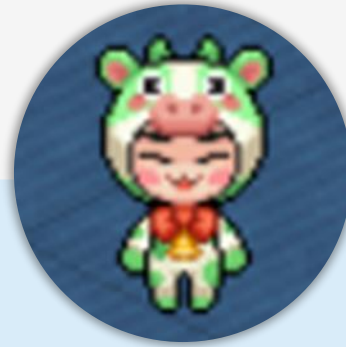
구민영

- 코드 분석 및 정리
- DB 스키마 설계
- Firebase 백엔드 DB 개발
- 프론트엔드 일부 기능 개발



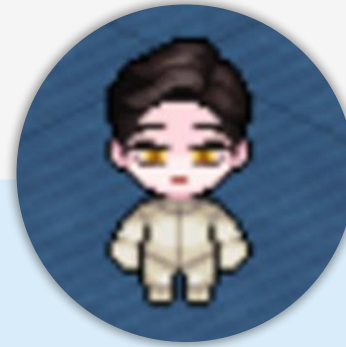
박찬진

- 프로젝트 기획 및 일정 총괄
- 코드 분석 및 정리
- 회의 내용 정리
- 모델링
- 키워드 추출 및 시각화



장다은

- 코드 분석 및 정리
- 모델링
- 콘텐츠 api 연결
- 키워드 강조한 요약문 생성



전우영

- 코드 분석 및 정리
- UI/UX 통합
- 플러터 프론트엔드 개발
- EC2 서버, 파이어베이스 백엔드 개발

3. 프로젝트 진행 프로세스

계획

11/22 ~ 11/29

- 📌 프로젝트 목표 정의
- 📌 기술 스택 선택

분석

11/27 ~ 12/01

- 📌 시장 조사
- 📌 비즈니스 모델 분석

설계

12/04 ~ 12/12

- 📌 DB 및 서버 구조 설계
- 📌 시스템 아키텍처 설계

개발

11/30 ~ 12/22

- 📌 노트 인식, 키워드 추출 및 시각화
- 📌 텍스트 요약, 콘텐츠 API 연결
- 📌 UI 구현
- 📌 DB 및 백엔드 구현

테스트

12/26 ~ 01/02

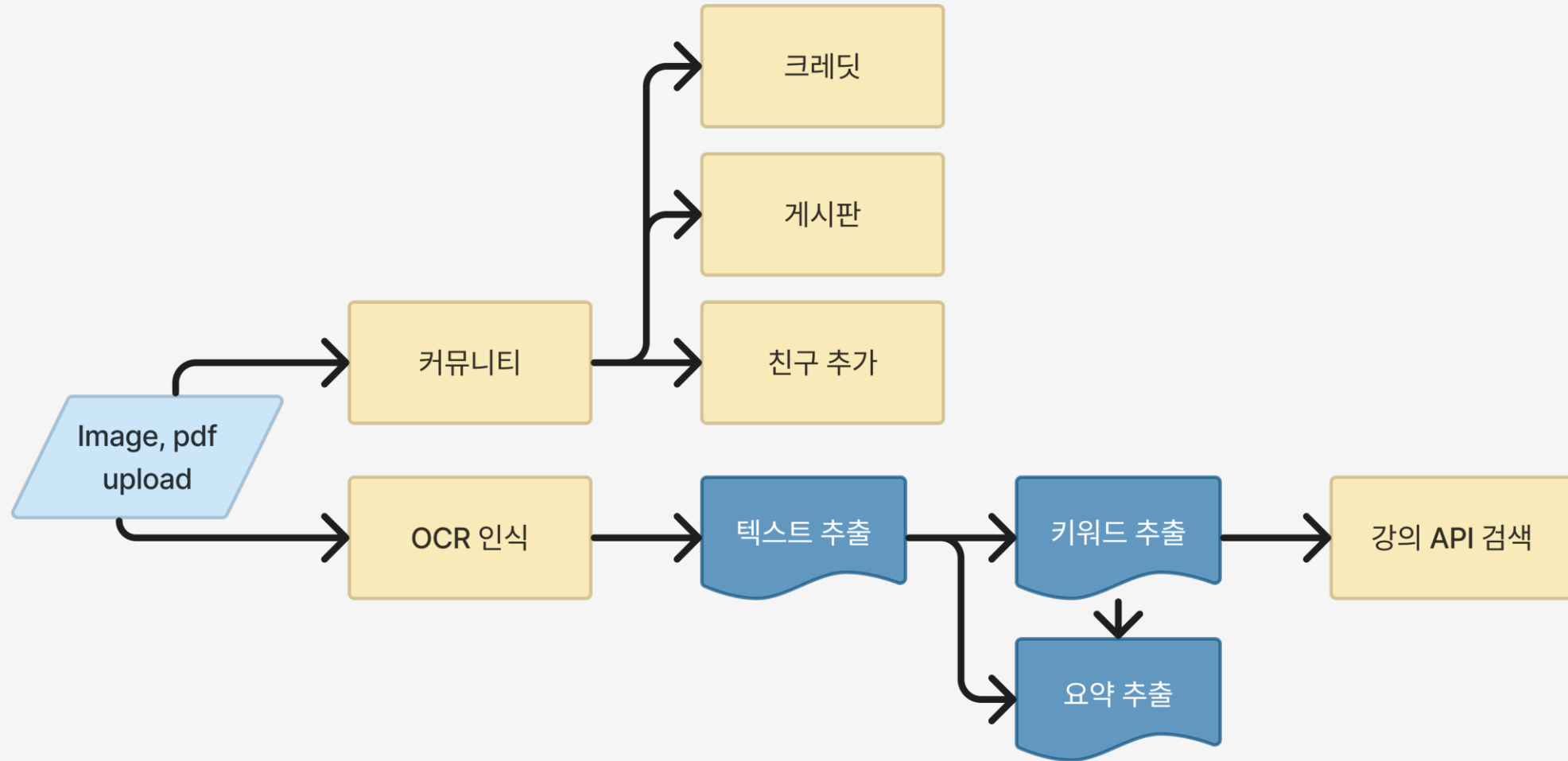
- 📌 유닛 테스트 및 통합 테스트
- 📌 버그 수정 및 성능 최적화

종료

01/02 ~ 01/04

- 📌 최종 문서 작성
- 📌 결과물 검토 및 평가

4. 프로젝트 결과



4. 프로젝트 결과



백엔드

- 📌 Firebase
- 📌 AWS EC2



키워드 추출

- 📌 KeyBERT
- 📌 TextRank 알고리즘



텍스트 요약

- 📌 KoBART

4. 프로젝트 결과 (백엔드)

번호	도메인 (1)	등록일	만료일
1703696	[수정] [삭제] wooyoung-project.kro.kr	2023-11-20	2024-02-20
[1]			

인스턴스 (1) 정보

Instance를 수정 또는 (case-sensitive) 태그로 찾기

인스턴스 ID

인스턴스 상태

인스턴스 유형

상태 검사

정보 상태

가용 영역

퍼블릭 IPv4 DNS

퍼블릭 IPv4

단일적 IP

wooyoung-pro...

i-04c73d7299dc0972

실행 중

t2.micro

2/2개 검사 통과

경보 없음

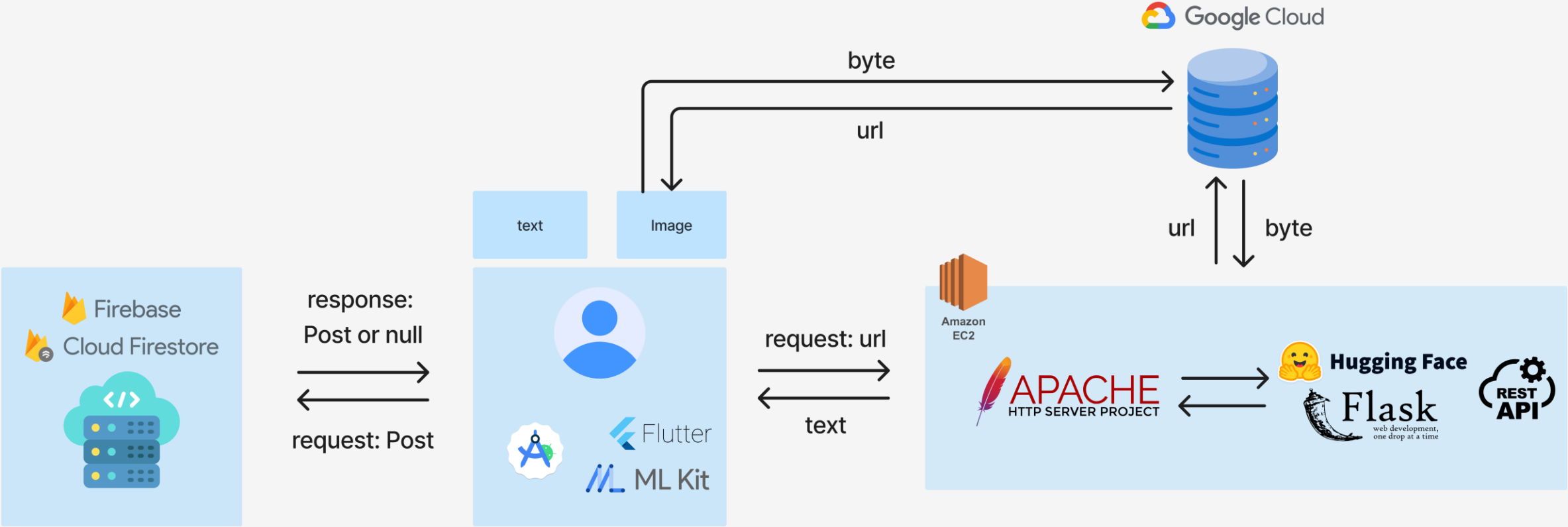
+

ap-northeast-2c

ec2-43-200-79-252.ap...

43.200.79.252

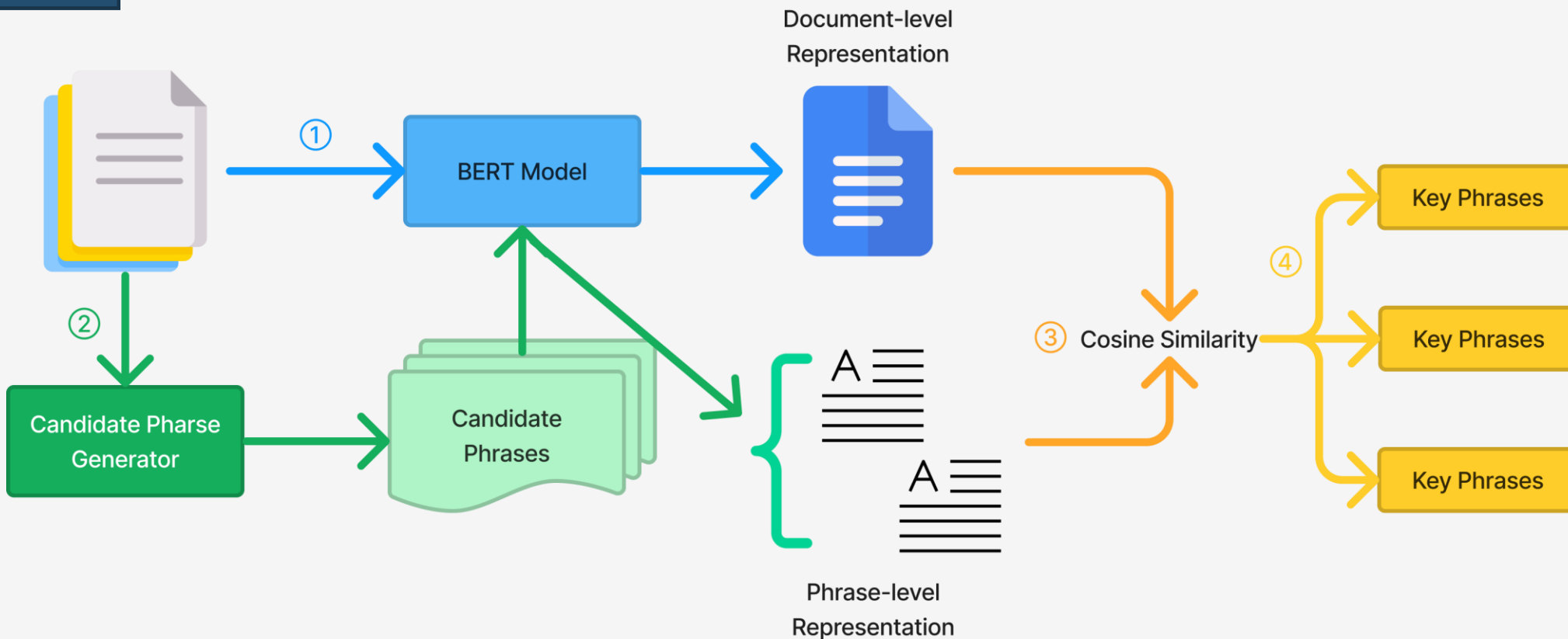
43.200.79.252



4. 프로젝트 결과 (키워드 추출) – 모델 개요



4. 프로젝트 결과 (키워드 추출) – 모델 분석



📌 Sentence-Transformers
📌 Gensim

📌 Flair
📌 USE

📌 Spacy

4. 프로젝트 결과 (키워드 추출) - 개선

Keyword: 04_diffusion_furnace 공정, Score: 0.5766
Keyword: 응용물리학과, Score: 0.4085
Keyword: 의존성, Score: 0.355
Keyword: dry와, Score: 0.3329
Keyword: 공정의, Score: 0.3232

Default 모델 (All-MiniLM-L6-v2)

Keyword: 산화막을, Score: 0.3244
Keyword: 산화막의, Score: 0.3151
Keyword: 산화막이, Score: 0.3084
Keyword: 산화막, Score: 0.3048
Keyword: dry에서, Score: 0.2739

distiluse-base-multilingual-cased-v1

- | | |
|---------|-------------|
| 1. 산화막 | 7. 계면 특성 |
| 2. Dry | 8. CMOS 반도체 |
| 3. Wet | 9. 실리콘 기판 |
| 4. 밀도차 | 10. 전류 방전 |
| 5. 성장속도 | 11. 특성 |
| 6. 밀도 | 12. 나노 단위 |

GPT 출력 결과 (텍스트 전체 입력)

→ TextRank 알고리즘 사용

총 10개의 pre-trained model 비교

그 중 GPT와 유사한 키워드를 추출하는 모델을 선택

TextRank만 사용:

막, 산화, 반도체, 계면,
가스, 속도, 산, 특성, 수,
균일, ...

distiluse-base-multilingual-cased-v1:
('밀도', 0.2361),
('반도체', 0.228),
('압력', 0.2219),
('산화', 0.192)

All-MiniLM-L6-v2 :
('공정', 0.4935),
('방전', 0.4728),
('특성', 0.4643),
('증착', 0.4263),
('평균', 0.4175)

- | | |
|--------|--------|
| 1. 반도체 | 10. 영향 |
| 2. 산화 | 11. 증착 |
| 3. 공정 | 12. 압력 |
| 4. 밀도 | 13. 산소 |
| 5. 열 | 14. 내부 |
| 6. 계면 | 15. 방향 |
| 7. 속도 | 16. 평균 |
| 8. 특성 | 17. 두께 |
| 9. 균일 | 18. 방전 |

4. 프로젝트 결과 (텍스트 요약) – 모델 개요

반도체 제조 기술의 이해

2 주차 (23. 9. 11)

단원: 04_Diffusion_Furnace공정

페이지수: 147-154

학과: 응용물리학과

이름: 박찬진

실리콘 산화막의 물성에서 Dry와 Wet 사이의 밀도차 이유와 이로 인한 영향?

건식산화는 순수한 산소를 사용하여 산화막을 만드는 방식으로, 성장속도가 느리지만 산화막의 두께는 얇고 밀도가 높습니다. 습식 산화는 산소와 용해도가 큰 수증기를 함께 사용하여기 때문에 성장속도가 빠른 대신 산화막의 두께가 두껍고 낮은 밀도가 특징입니다.

Dry에서 더 우수한 계면 특성. 계면 특성?

-> 계면에서의 분자 사이의 상호 작용과 관련된 물리적 성질.

산화막 계면 특성이란 CMOS반도체의 실리콘 기판과 산화막이 맞닿는 경계면에서 전류가 방전되는 결함 등 특성을 가리킨다. 반도체의 크기가 작아지면 CMOS 소자 내 산화막의 두께도 얇아지는데 산화막이 얇을 수록 방전이 많이 일어나는 문제점이 발생한다. 특히 나노 단위의 고집적 반도체로 갈수록 이러한 계면 결함을 줄이는 것이 반도체 수율(생산 성공율)을 높이는 관건이 된다.

Thermal Conductivity: 열 전도율 / Index of Refraction: 굴절율

Remider: Stress? 외력에 저항하기 위해 내부에서 생기는 힘.

산화의 Si 결정 방향 의존성? 0918 정리 참고

균일 분해? 공유 결합이 끊어질 때, 결합을 구성하고 있는 한 쌍의 전자가 각각 한 개씩 분리되는 현상.

Diffusion: LPCVD 주로 사용 (Low Pressure Chemical Vapor Deposition)

-> 다른 CVD들에 비해 Step Coverage가 우수

튜브 내부 압력이 저압인 상태에서 가스 유입 -> 쉽게 표면으로 이동, 흡착 -> 열 에너지에 의해 반응 생성물 형성 -> 나머지 가스 탈착 후 배기, 생성물들 축적 - 박막 형성

Deposition 주요 인자: Quality, Thickness Uniformity, Filling, Step Coverage, Aspect Ratio

균일도: 후공정에 영향 많이 미침 / 스텝 커버리지: 수직, 수평 방향 간 증착 비율의 균일도 / 종횡비: 높이와 폭의 비율 (=단차), 고종횡비: 폭이 좁음, Filling이나 Step coverage가 좋지 않아 공정의 어려움

증착 속도에 영향을 주는 인자: 온도, 압력, 가스량, 웨이퍼의 수량

증착 속도 \propto 온도, 압력, 가스량 \propto 1/(웨이퍼 수)

평균 자유 행로: 충돌 직후부터 다른 입자와의 충돌 직전까지의 평균 이동 거리

1



GPT-2

반도체 제조 기술의 이해 2주차(23.9.11)단원:04_Diffusion_Furnace공정 페이지수: 147-154
학과: 응용물리학과 이름: 박찬진 실리콘 산화막의 물성에서
...
(요약 실패)

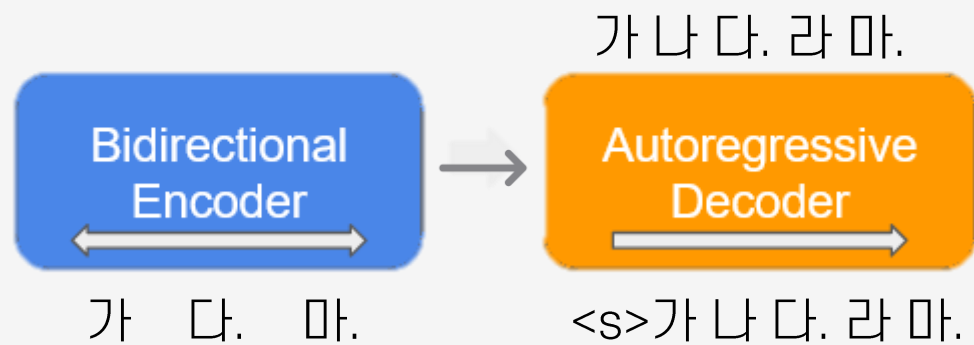
BERT

Dry Dry's name means 'dry' and 'wet' in English. Dry means "to dry" or 'to remove' Wet
(요약 실패)

KoBART

CMOS 반도체의 실리콘 기판과 산화막이 맞닿는 경계면에서 전류가 방전되는 결함 등 특성을 가리키는 산화막 계면 특성이란 CMOS 반도체의 실리콘 기판과 산화막이 맞닿는 경계면에서 전류가 방전 되는 결함 등 특성을 가리킨다.

BART, KoBART



- Autoencoder
- 노이즈된 입력 복구 학습 방식

4. 프로젝트 결과 (텍스트 요약) - 개선

KoBART

GPT-3: 반도체 제조 과정에서 건식 산화는 높은 밀도의 얇은 산화막을 생성하며 우수한 계면 특성을 제공하는 반면, 습식 산화는 빠른 성장속도로 두꺼운 산화막을 형성하지만 낮은 밀도를 가진다.

KoBART: CMOS 반도체의 실리콘 기판과 산화막이 맞닿는 경계면에서 전류가 방전되는 결함 등 특성을 가리키는 산화막 계면 특성이란 CMOS 반도체의 실리콘 기판과 산화막이 맞닿는 경계면에서 전류가 방전 되는 결함 등 특성을 가리킨다.

	ROUGE-1	ROUGE-2	ROUGE-L
Precision	0.086956	0.0	0.086956
Recall	0.117647	0.0	0.117647
F1	0.099999	0.0	0.099999

	BERT SCORE
Precision	0.648915
Recall	0.644450
F1	0.646675

4. 프로젝트 결과 (텍스트 요약) - 개선

KoBART

GPT-3: 반도체 제조 과정에서 건식 산화는 높은 밀도의 얇은 산화막을 생성하며 우수한 계면 특성을 제공하는 반면, 습식 산화는 빠른 성장속도로 두꺼운 산화막을 형성하지만 낮은 밀도를 가진다.

KoBART: 건식산화는 순수한 산소를 사용해 산화막을 만드는 방식으로 성장속도가 느리지만 산화막의 두께는 얇고 밀도가 높다 습식 산화는 산소와 용해도가 큰 수증기를 함께 사용하기 때문에 성장속도가 빠른 대신 산화막의 두께가 두껍고 낮은 밀도가 특징이다

	ROUGE-1	ROUGE-2	ROUGE-L
Precision	0.217391	0.041666	0.217391
Recall	0.172413	0.032258	0.172413
F1	0.192307	0.036363	0.192307

	BERT SCORE
Precision	0.761863
Recall	0.772249
F1	0.767021

4. 프로젝트 결과 (텍스트 요약) - 개선

KoBART + Keyword 강조

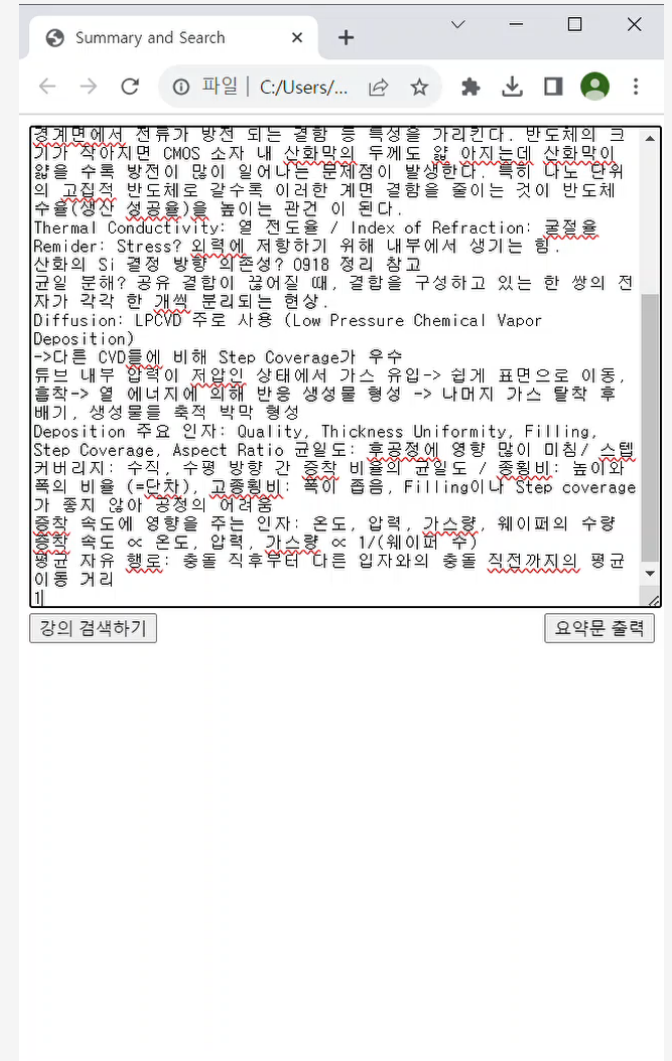
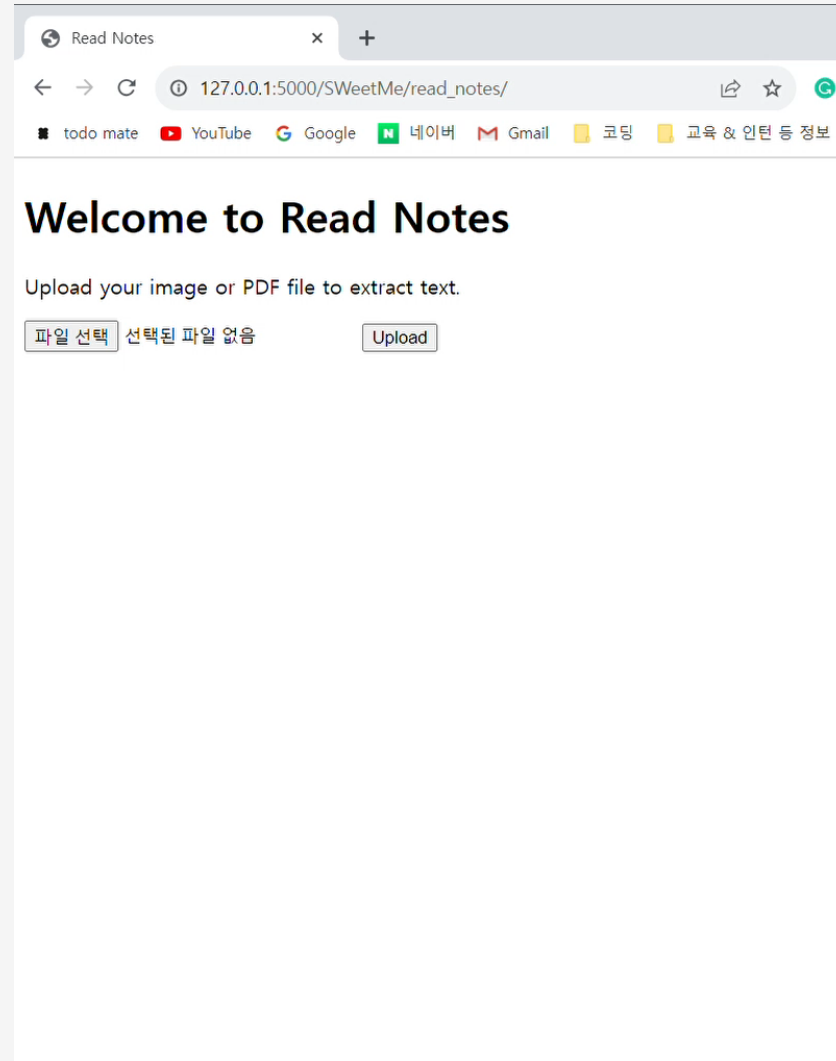
GPT-3: 반도체 제조 과정에서 건식 산화는 높은 밀도의 얇은 산화막을 생성하며 우수한 계면 특성을 제공하는 반면, 습식 산화는 빠른 성장속도로 두꺼운 산화막을 형성하지만 낮은 밀도를 가진다.

KoBART: 실리콘 산화막의 물성에서 Dry와 Wet 사이의 밀도차 이유와 이로 인한 영향은 CMOS반도체 기판과의 경계면에서 전류가 방전 되는 결함 등 특성을 가리킨다.

	ROUGE-1	ROUGE-2	ROUGE-L
Precision	0.043478	0.0	0.043478
Recall	0.047619	0.0	0.047619
F1	0.045454	0.0	0.045454

	BERT SCORE
Precision	0.693337
Recall	0.693756
F1	0.693548

4. 프로젝트 결과 (시연 영상)



5. 자체 평가 및 보완



구민영

처음으로 어플리케이션을 만들어보는 것이 흥미로웠음. 특히 플러터로 어플리케이션을 만들어보는 것은 아주 좋은 경험이 되었음. 다만 데이터베이스를 MySQL 같은 관계형 데이터베이스가 아닌, NoSQL 데이터베이스(파이어베이스)를 사용하였기에 복잡한 쿼리를 구현하는 것이 어려웠음. (만약 다음에 유사한 프로젝트를 진행한다면 MySQL로 프로젝트를 진행해보고 싶음. 또한 프론트엔드 부분의 UI를 기타 SNS 어플처럼 좀 더 꾸며볼 기회가 있으면 좋겠음.

초반에 조사 및 분석을 더 꼼꼼하게 했더라면 개발에 필요한 것들과 구현이 예상대로 진행이 가능했을 것이라 생각함. 스스로 기획이 탄탄하지 못했다는 점에서 아쉬움이 남지만 이후 키워드 추출에서 AI 모델을 다른 알고리즘과 함께 사용해보며 유의미한 키워드 추출이 되는 것에서 성과가 있었다고 생각함.



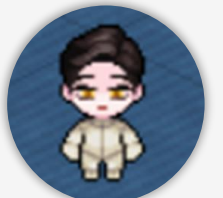
박찬진



장다은

요약 작업에 특화된 모델을 선정하고 최적화 하는 과정에서 추출된 키워드를 강조하는 알고리즘을 추가해 실제로 성능면에서 개선을 보였다는 점이 성과로 생각됨. 또한, 논문 요약 데이터를 활용하여 전이학습 시켰을 때 의미있게 성능이 올라갔다는 점에서 추후 정확도를 높인 필기 요약 제공이 기대됨. 모듈화 한 채로 개발하다 보니 통합테스트에서 발생할 문제들을 예상하지 못함. 모델을 서빙하는 것이 호환성 문제가 많았고 통합 테스트 환경 고려가 필요함.

개인적으로 플러터라는 새로운 프레임워크를 사용해서 서비스를 만드는 경험이 재밌었음. 상태 관리 같은 고급 활용 기술들을 추가적으로 학습하여 보완하면 좋을 것 같음. EC2 프리 티어로는 서비스를 모두 돌릴 수 없어서 아쉬웠음. 기회가 된다면 좀 더 좋은 스펙으로 서버를 돌리면 좋을 것 같음.



전우영

Q & A

감사합니다.