

TEAM HOTSIX

기술 기반 검색 서비스 ‘search me in tech’

2023 KPMG 아이디어톤

KPMG



윤희승 모운호 심소현 양정안 이상민 정세희

Contents

01 동기

1.1 제안 배경

02 주요 기능 소개

2.1 서비스 소개

2.2 아키텍처 디자인

03 시연 영상

04 시스템 및 기술 소개

4.1 기술 키워드 기반 기업 검색

4.2 기업 키워드 기반 유사 기업 추천

4.3 기업 키워드 기반 최근 뉴스 기사 감성 분석

4.4 기업 키워드 기반 재무제표 요약

05 결론 및 제언

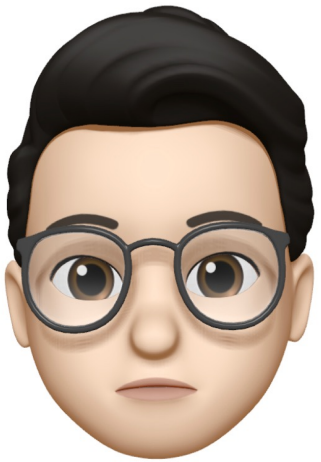
5.1 결론

5.2 제언 및 추후 계획

5.3 기대 효과

06 Appendix

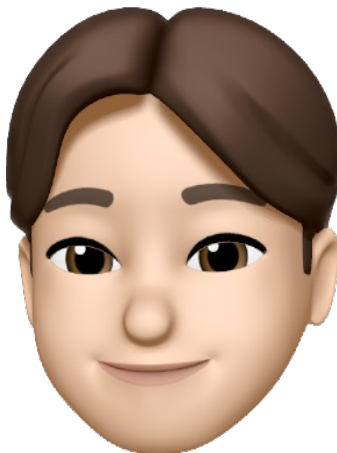
우리는 기업을 위해 기술을
기반하여 기업을 찾아주는
시스템을 개발한
HOTSIX 팀이야 -!



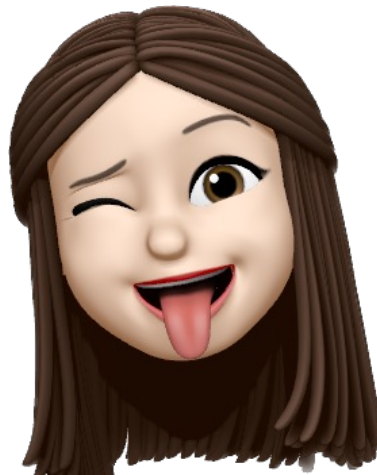
윤희승



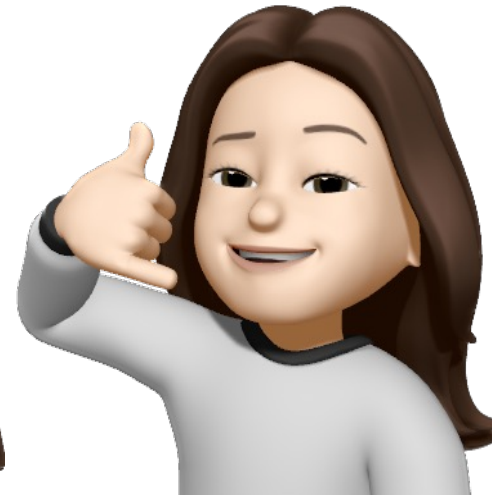
이상민



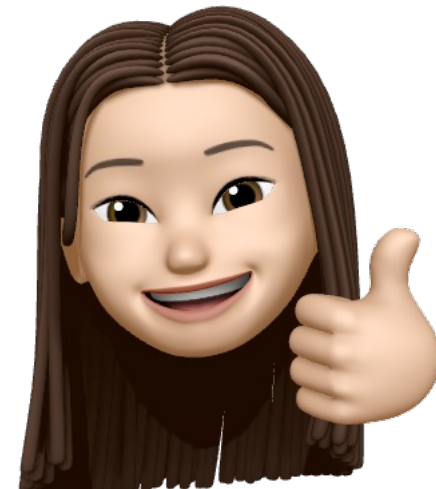
모운호



정세희



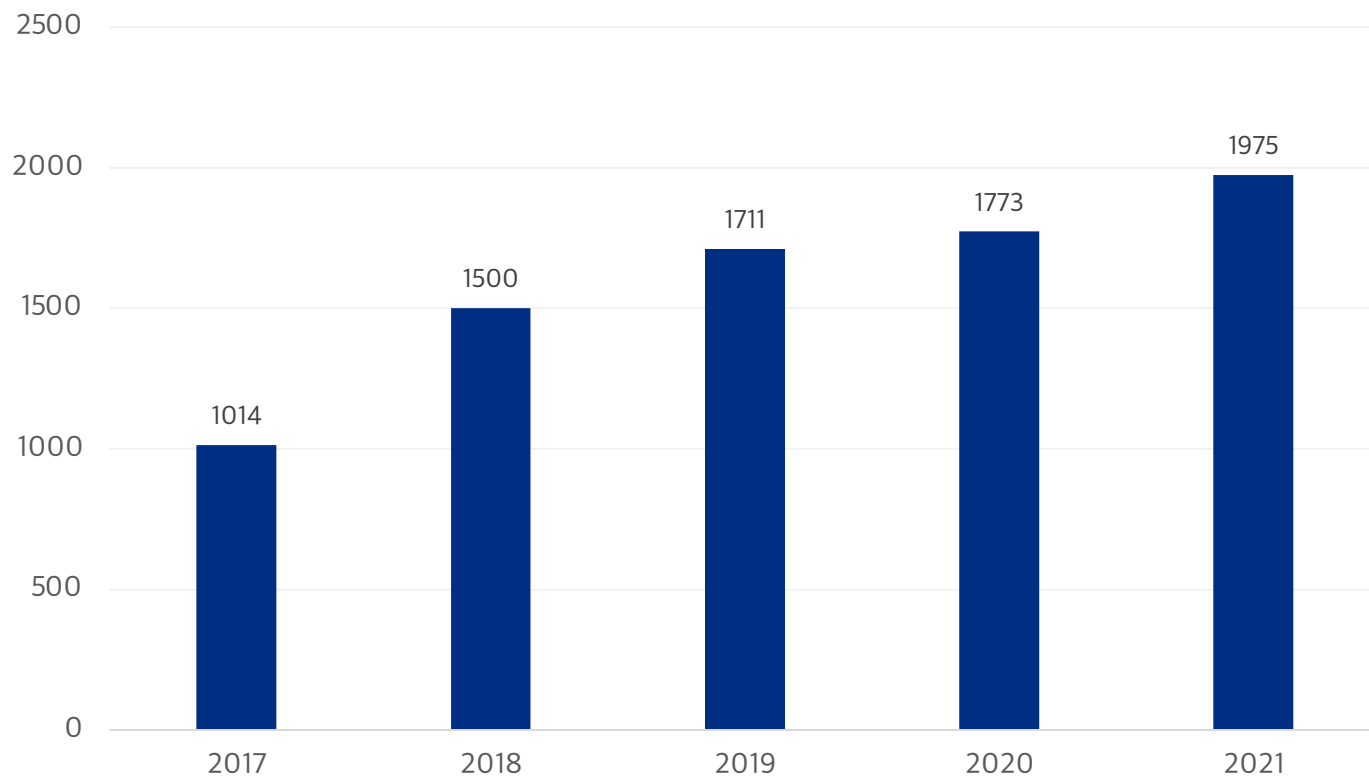
양정안



심소현

‘ 기술 기반 기업 수 증가 ’

연도별 4차산업 기술 활용 기업 수



통계청 기업활동조사 2017-2021 에 따르면, 연도별 4차 산업 기술 활용 기업 수에서 연도가 올라갈 수록 기술 발전의 가속화와 기술 중심 기업들의 등장으로, 산업 전반에서 기술 트렌드 변화에 빠르게 대응하려는 움직임이 활발함

이에 따라, 기술에 투자하고자 하는 기업들과 기술을 가진 기업 간의 매칭 수요가 증가하고 있음

기업의 기술 정보 파악의 문제

기술을 보유하고 있는지,
연관성만 있는지 알 수 없는 경우

검색어가 정확히 일치해야 하여,
인접한 키워드로는 검색이
되지 않는 경우

검색이 아닌 필터링으로만
관련 기업 조회가 가능한 경우

자율주행로봇



AA 기업
자율주행로봇 개발 중인 BB사에 투자 유치...

BB 기업
Bb 대표, 자율주행로봇 개발에 대한 비전 제시...

CC 기업
자율주행로봇 개발 기업 DD사 인수 추진...

불량 검출



AA 기업
제조 공정에서의 불량 검출 자동화 솔루션 제공

BB 기업
공정에서 불량을 자동으로 검출하는 솔루션 개발

CC 기업
제조 공정에서의 불량을 자동으로 탐지하는 솔루션

범주 1

...



범주 2

...



범주 3

...



항목 1
항목 2
항목 3
...

✓ 기술 키워드 기반 기업 정보 제공

챗봇 기반 기술 키워드를 통해 관련 기업과
해당 기업의 주요 정보 제공

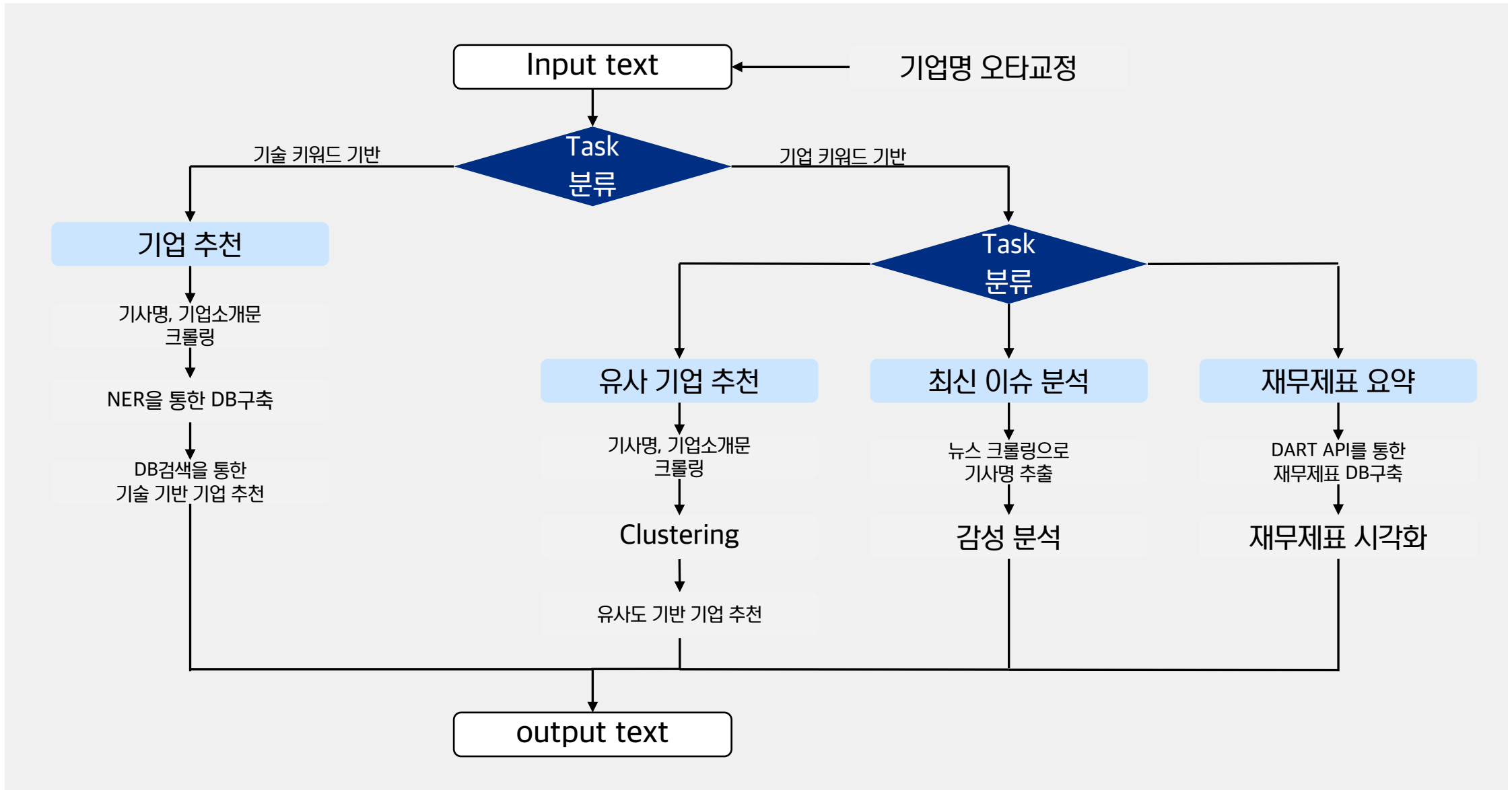
“우리 분야에 특화된 음성인식 기술을 도입하고 싶은데...”

OO 분야에서 음성을 자동으로 텍스트화
하는 기술을 가진 기업을 찾고 싶어.

AA기업에 대한
정보를 보여드리겠습니다.

AA 기업

입력된 키워드에 따라 관련 기업과 기업 정보 제공





편집 기반 알고리즘(Levenshtein(Edit) distance)

입력된 텍스트를 기업명 DB의 올바른 기업명 중
Levenshtein distance 기준으로 가장 가까운 단어로 치환

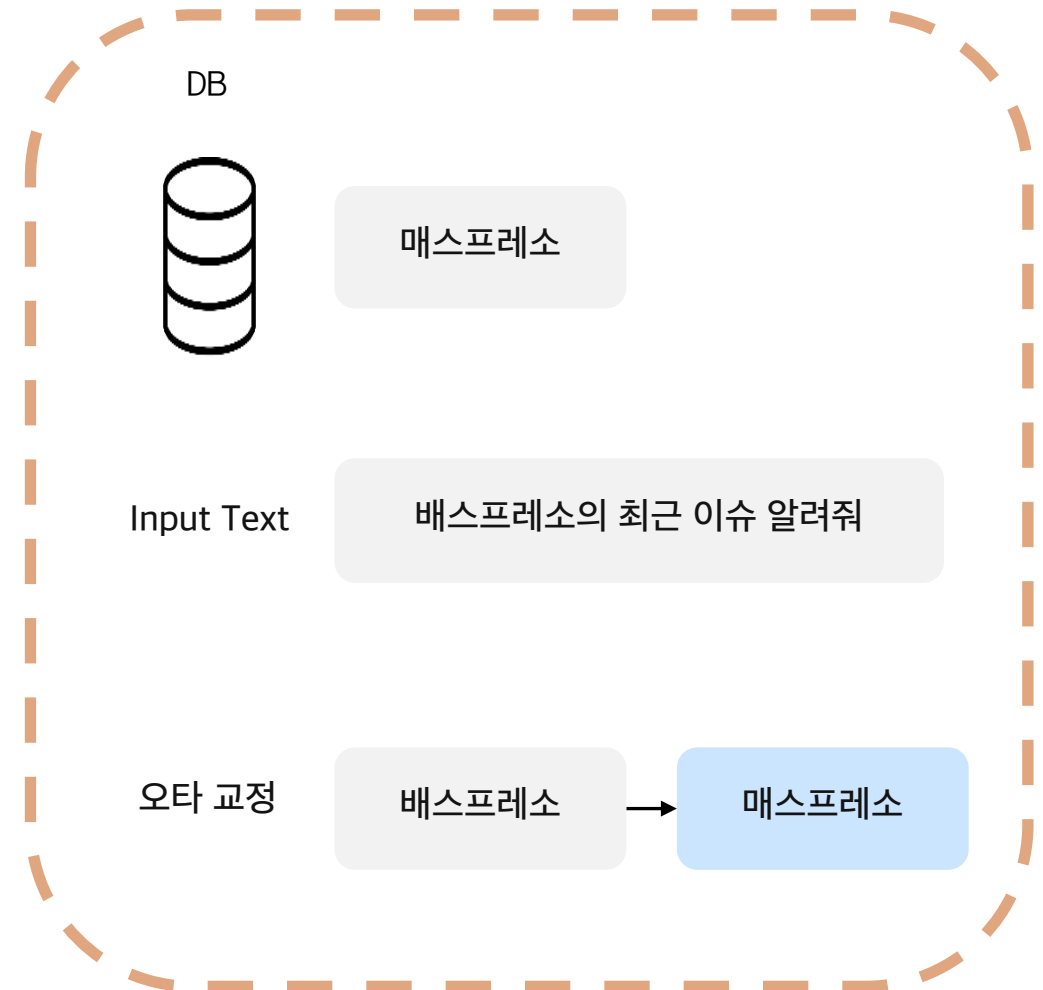
String distance

단어, 문장과 같은 text간의 형태적 유사성을 정의 하는 방법

Levenshtein(Edit) distance

string distance의 대표적인 Metric

text1를 text2로 변환하기 위한 최소 횟수를 두 text의 거리로 정의



기업 소개문 및 뉴스에서 annotation을 통해 기술 키워드 추출하여 기업별 기술 키워드 데이터베이스 구축



기업 공시, 관련 뉴스로부터
기업의 원천 기술 데이터 확보

annotation을 통해
기술 키워드 추출

추출한 정보를 기반으로
기업명-기술 키워드 DB 구축

입력된 기술 키워드로부터, 해당 기술을 보유한 기업을 검색 결과로 제시

기업 검색

OO 분야에 특화된 챗봇 솔루션을
제공하는 기업이 있을까?

OO 분야

챗봇

...

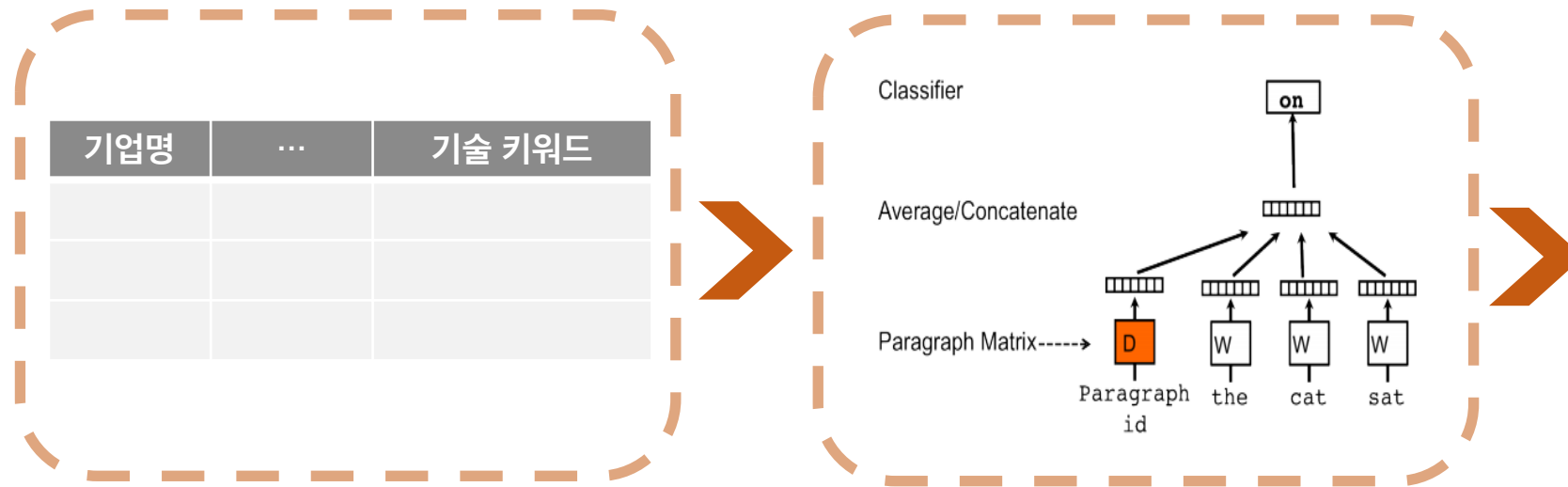
기술 키워드가 포함된
질문 입력

텍스트 전처리를 거쳐
문장 내 기술 키워드 추출



DB에서 기술 키워드를 조회하여
해당하는 기업 출력

입력된 기업과 유사한 기술을 가진 기업을 검색 결과로 제시
클러스터링을 활용하여 기업 간 유사도 비교

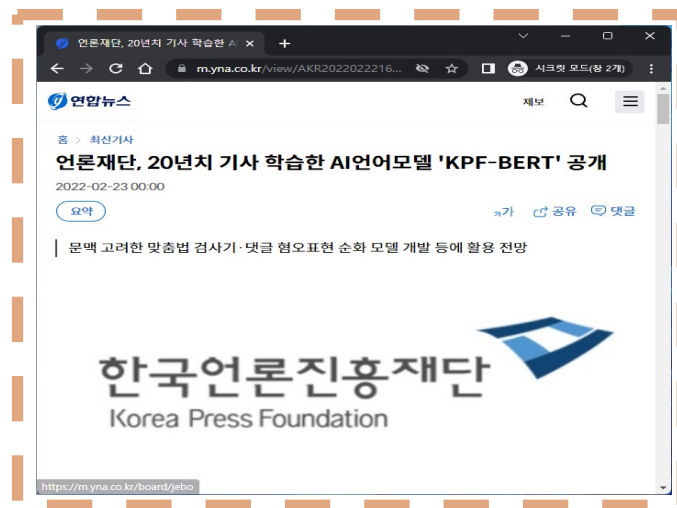
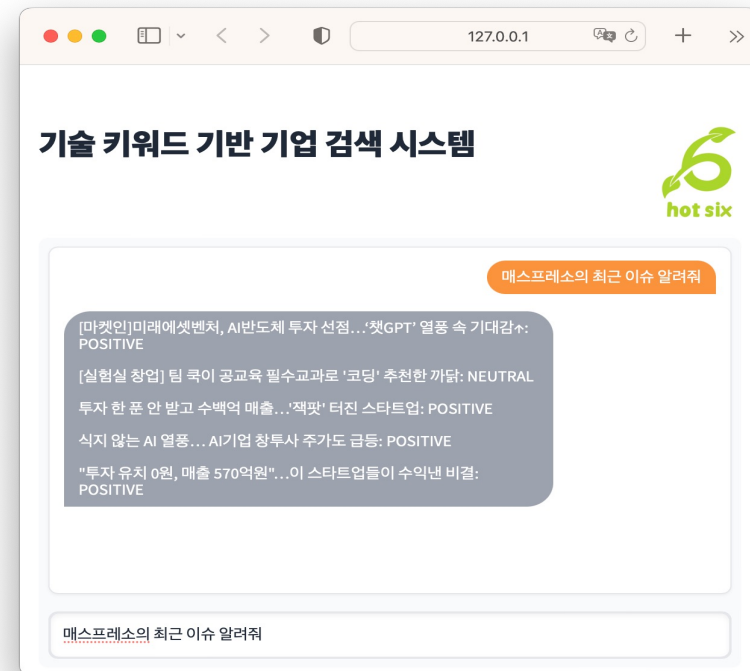


Annotation으로 추출한 정보로
기업명-기술 키워드 DB 구축

기술 키워드를 추출하여
Doc2Vec으로 유사도 계산

유사도가 가장 높은
기업 출력

입력된 기업의 최근 기사 추출 후 감성 분석하여 검색 결과로 제시

뉴스 크롤링으로 최근 기사
제목 수집Korfin 뉴스기사 감성 분석 데이터로
KPF-BERT fine-tuning검색한 기업에 대한 기사제목에 대한
감성분석 결과 출력

감성분석 모델 학습 과정

Finetuning 대상의 모델은 여러 모델을 검토한 결과, KPF-BERT를 선택
→ 뉴스 데이터를 기반으로 모델 학습을 진행하였기 때문

데이터셋의 경우 기사명과 해당 기업, 기사의 감정이 분류된 Korfin 데이터셋을 이용

ex)

기사명: "삼성물산은 최근 삼성전자 등 계열사의 공격적인 설비투자 계획에 힘입어, 이미 상반기에만 1.4조원(해외 미포함)의 그룹공사를 수주했으며, 2010년 연간으로는 약 2.5조원의 그룹공사 수주가 예상된다."

기업명: "삼성물산"

감정분석: "POSITIVE"

기사명: "게다가 KT&G의 국내 MS는 1분기 대비 1.4%pt 하락한 58.8%로 의 추정치 59.4%를 하회했다."

기업명: "KT&G"

감정분석: "NEGATIVE"

입력된 기업명으로부터 공시된 재무제표를 추출 후 요약하여 검색 결과로 제시

벤처확인기업 공시

벤처확인기업 공시제도 안내

벤처기업확인기관은 「벤처기업육성에 관한 특별조치법」 시행령 제18조의5(벤처기업에 관한 정보의 공개)에 따라서 벤처확인기업에 대해서 아래 정보를 공시하고 있습니다.

- 일반정보: 상호, 업종, 등기부상의 법인등록번호, 주소, 전화번호, 주요 제품 및 그 변경사항
- 재무정보: 대차대조표와 손익계산서
- 투자 관련 정보: 투자받은 금액, 투자시기 및 그 변경사항(벤처투자유형에 한함)
- 벤처기업확인사: 발급일, 유효기간 및 그 변경사항

※ 공시내용이 사실과 다를 경우 벤처기업확인기관 고객센터(1566-6487)로 연락하여 주시고, 본 공시정보를 활용하여 행한 결정과 그 결과에 대해서 벤처기업확인기관은 어떠한 법적 책임도 지지 않음을 알려 드립니다.

기업명 대표자명 사업자번호

순수 기업명을 입력해주세요. (주)한국



구 분	재무비율	산 식
안정성	유동비율	(유동자산÷유동부채)×100
	자기자본비율	(자본총계÷부채와자본총계)×100
	부채비율	(부채총계÷자본총계)×100
수익성	총자본세전계속사업이익률	(세전계속사업이익÷부채와자본총계)×100
	매출액세전계속사업이익률	(세전계속사업이익÷매출액)×100
	금융비용부담률	(이자비용÷매출액)×100
활동성	총자산회전율	매출액÷부채와자본총계
	운전자금회전기간	(매출채권+미수금-매입채무)÷월매출액
	재고자산회전기간	재고자산÷월매출액
생산성	부가가치율	(부가가치÷매출액)×100
	설비투자효율	(부가가치÷(유형자산-건설중인자산))×100
	총자본투자효율	(부가가치÷전기 및 당기 평균 총자산)×100
성장성	총자산증가율	((당기 총자산-전기 총자산)÷전기 총자산)×100
	매출액증가율	((당기 매출액-전기 매출액)÷전기 매출액)×100
	순이익증가율	((당기 순이익-전기 순이익)÷전기 순이익)×100

※ 신규 재무제표가 수집되어, 이전 공시자료의 동일한 결산년도의 재무비율 값과 달라질 수 있음.



기술 키워드 기반 기업 검색 시스템

매스프레소의 재무제표 알려줘

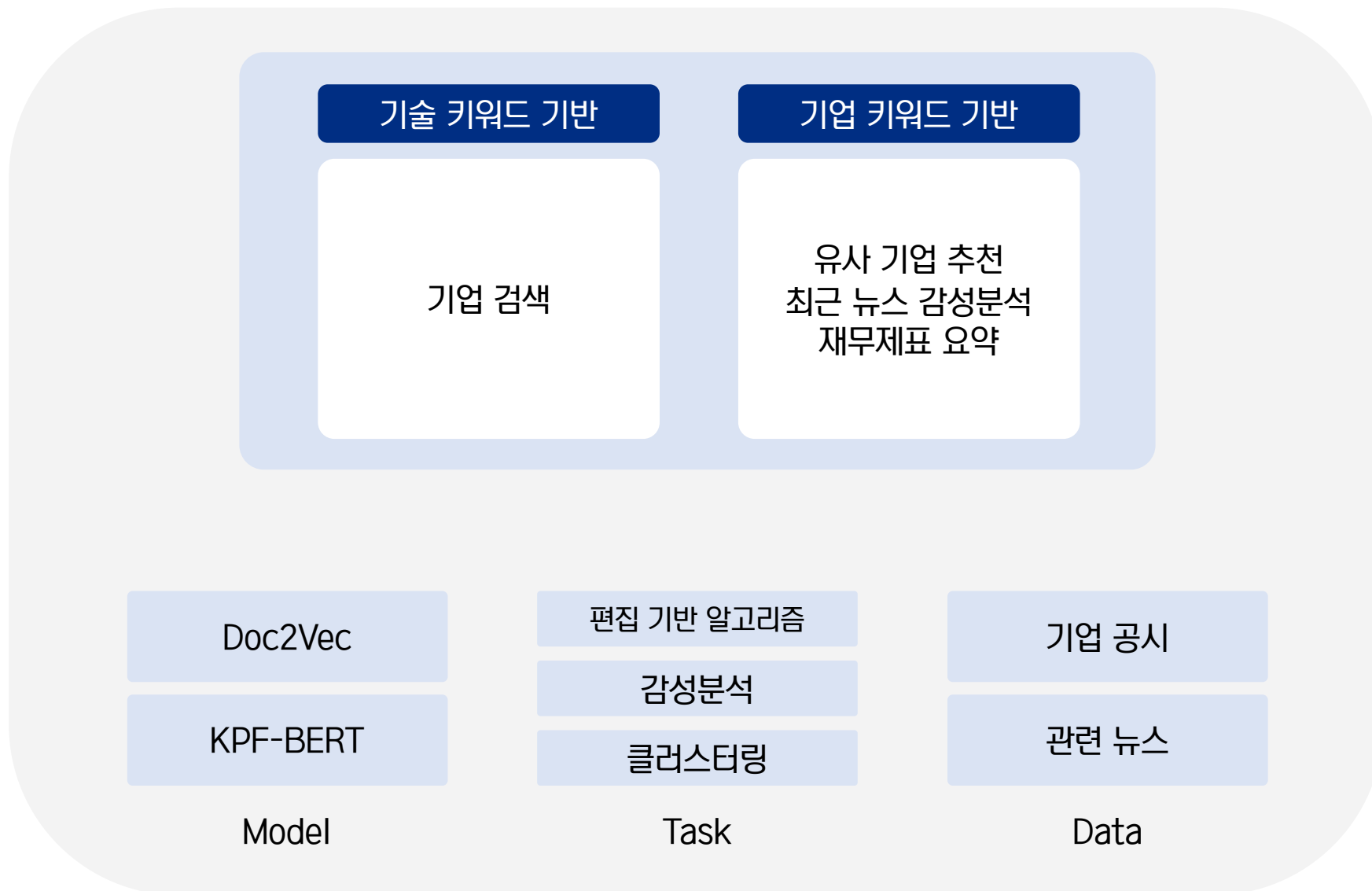
유동비율: 1641.445115
자기자본비율: 93.52725626
부채비율: 6.920703116
총자산회전율: 2.6970988
총자산증가율: 110.9956169
매출액증가율: 267.271494
순이익증가율: 114.6649188
(기준년도: 2021)

매스프레소의 재무제표 알려줘

벤처확인기업 공시 시스템을 통해
기업정보에서 재무제표 추출

신용보증기금 자료 기준
7개 주요 지표 산출

재무제표 출력



1. 기술 키워드를 기반으로 관련성이 높은 기업을 추천하는 시스템을 제안하였음
2. 기업의 기술 키워드는 기업 관련 기사와 기업 설명문을 토대로 팀원들이 annotation하여 추출했고, 이를 바탕으로 기업-기술 키워드 DB를 구축하였음
3. 해당 DB를 활용하여 입력된 기술 키워드와 관련성이 높은 기업을 조회하고, Doc2Vec 모델로 유사한 기업 간 클러스터링을 진행하였음
4. 기업의 최신 기사 제목으로 감성분석을 진행하여 기업 이슈 분석을 진행하였음
감성분석은 한국언론진흥재단이 공개한 KPF-BERT를 fine-tuning하여 진행하였음
5. 재무 정보는 중소기업벤처부의 벤처기업확인 시스템을 통해 기업별 재무제표를 수집한 후, 신용보증기금에서 공개하는 지표를 기준으로 주요 재무 지표를 산출하였음

input: 엔씨가 연구개발하고 있는 ▲게임 ▲음성인식·화성 ▲컴퓨터 비전 ▲자연어 처리 ▲지식 공학 등 5가지 영역의 AI 기술에 대하여 총 24개의 세션을 열어 심도 있는 논의를 진행했다.

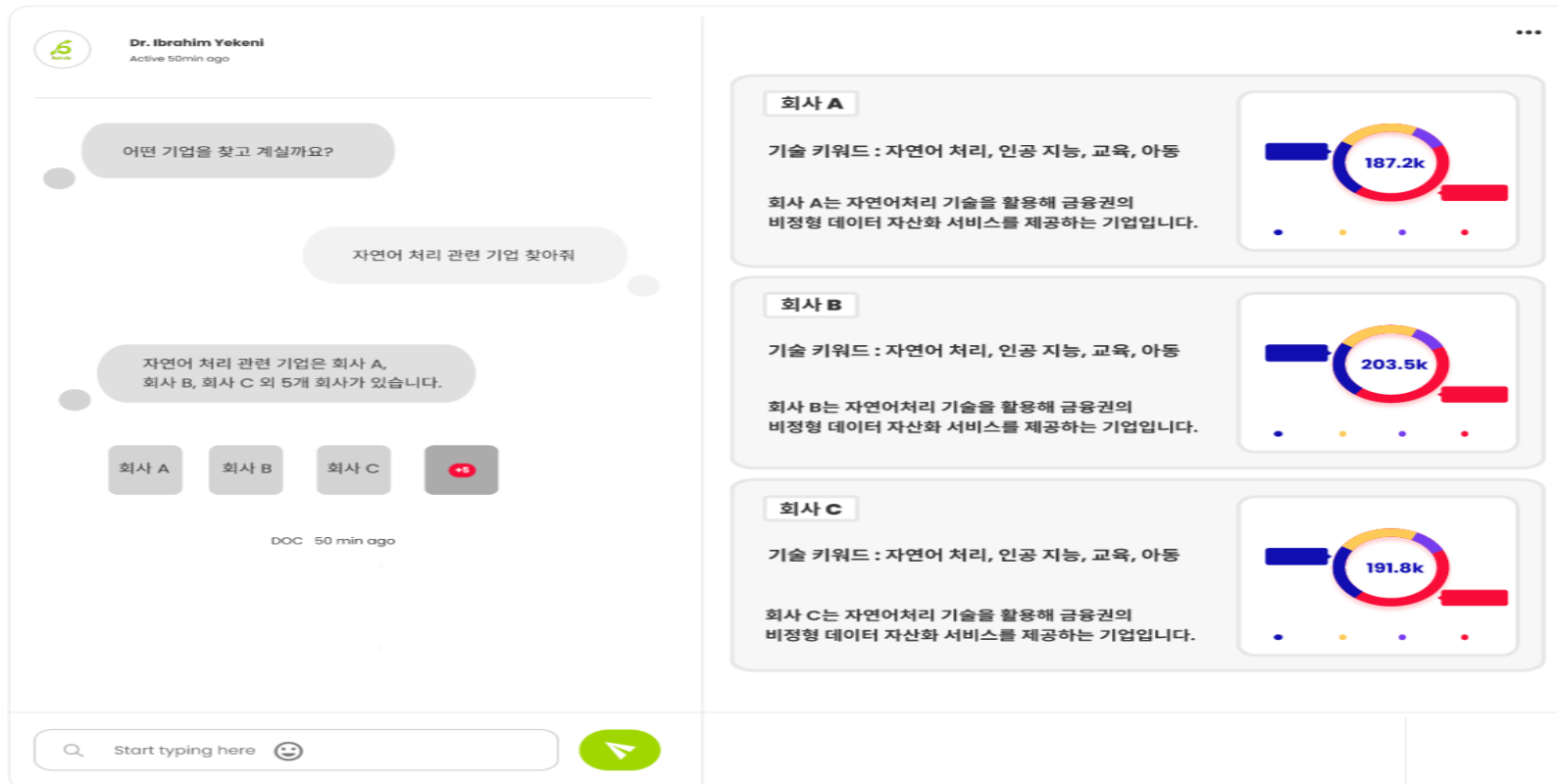
[illegible]

input: 고려대 AI대학원의 핵심 연구 분야는 딥러닝, 컴퓨터비전, 자연어처리, 음성인식, 빅데이터, 신경망 등과 'AI+X(헬스케어, 금융, 지능형에이전트, 게임, 자율주행, 국방 등)' 다.

[illegible]

시각화 기능 보완

재무 지표, 기업의 최신 뉴스 등을 우측 정보란에 시각화하여 제시하고자 하였으나 구현의 어려움.
향후 보완을 통해 우측 정보란으로 챗봇 결과를 출력하는 기술을 구현할 예정



The screenshot displays a chatbot interface. On the left, a chat window shows a conversation with 'Dr. Ibrahim Yekeni'. The user asks, '어떤 기업을 찾고 계실까요?' and the chatbot responds, '자연어 처리 관련 기업 찾아줘'. The chatbot then lists '회사 A', '회사 B', and '회사 C' as relevant companies. Below the chat window, there is a search bar with the text 'Start typing here' and a green arrow button.

On the right, a sidebar displays detailed information for three companies:

- 회사 A**: 기술 키워드: 자연어 처리, 인공 지능, 교육, 아동. 회사 A는 자연어처리 기술을 활용해 금융권의 비정형 데이터 자산화 서비스를 제공하는 기업입니다. Visualized value: 187.2k.
- 회사 B**: 기술 키워드: 자연어 처리, 인공 지능, 교육, 아동. 회사 B는 자연어처리 기술을 활용해 금융권의 비정형 데이터 자산화 서비스를 제공하는 기업입니다. Visualized value: 203.5k.
- 회사 C**: 기술 키워드: 자연어 처리, 인공 지능, 교육, 아동. 회사 C는 자연어처리 기술을 활용해 금융권의 비정형 데이터 자산화 서비스를 제공하는 기업입니다. Visualized value: 191.8k.

Each company profile includes a circular progress indicator and a bar chart.

기업 간 기술 교류 및 협업 준비 과정의 간소화

- 기업 매칭에 드는 시간과 비용 부담을 줄임으로써
사업 추진에 필요한 과정 전체를 단축시키는 효과 기대

산업 전반에서의 최신 기술 도입 활성화

- 기업 매칭에 대한 부담 감소로 기술 교류를 위한 기업 탐색이 활발해지고,
기업들이 보다 적극적으로 기술 교류에 나설 것으로 기대

기술 기반 기업 검색 AI의 발전과 확대 가능성

- 검색 알고리즘의 개선과 기술 및 기업 정보에 관한 추가 기능 도입으로,
기술 기반 검색 AI의 기술적 확장 가능성 전망

기술 키워드 기반 기업 검색 시스템

Thank you

유동비율: 1641.445115

자기자본비율: 93.52725625

부채비율: 6.920703116

총자산회전율: 2

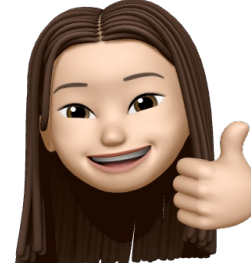
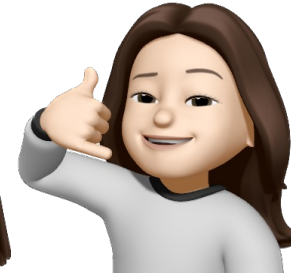
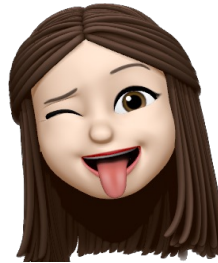
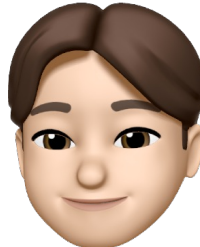
총자산증가율: 1

매출액증가율: 2

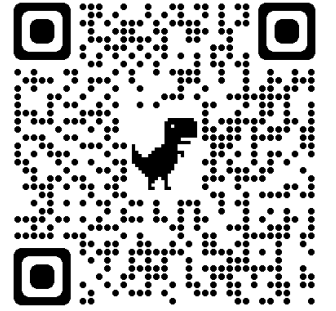
순이익증가율: 1

(기준년도: 2021)

KPMG 아이디어톤 | 윤희승 모운호 심소현 양정안 이상민 정세희



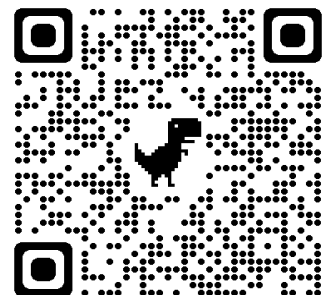
Github link



Demo-video link



Document link



매스프레소의 재무제표 알려줘

네이버 뉴스 API: <https://developers.naver.com/docs/serviceapi/search/news/news.md>

Korfin dataset: <https://huggingface.co/datasets/amphora/korfin-asc>

KPF-BERT: <https://github.com/KPFBERT/kpfbert>

KPF-BERT 관련 기사: <https://www.yna.co.kr/view/AKR20220222161500005>

중소기업벤처부 벤처확인기업 공사: <https://www.smes.go.kr/venturein/pbntc/searchVntrCmp>

신용보증기금 재무비율 자료:

<https://www.kodit.co.kr/kodit/na/ntt/selectNttList.do?mi=2806&bbsId=1002&ps=427>

국립국어원 모두의 말뭉치: <https://corpus.korean.go.kr/request/reasetMain.do#down>