

News Summary APP Using MachineLearning



NewSum

당신의 입맛을 맞춘 뉴스 어플리케이션 NewSum!

NewSum은 뉴스 요약 및 추천 서비스를 제공하는 AI 뉴스 어플리케이션입니다.
Machine Learning을 이용하여 개인화 맞춤 추천 서비스,
뉴스 요약본 및 헤드라인 생성 기능을 제공합니다.



Scan me

Team MUD

201601364 박주영(T) 201503665 홍승환 201600599 김아연 201701138 김혜원 201700124 이산가 비두샤

NewSum이 왜 필요해?

오늘 하루 일들을 단 **20가지**로! 하루 뉴스 브리핑!

하나하나 뉴스기사 읽지 말고 비슷한 뉴스를 묶어 **큰 뉴스** 만

오늘의 뉴스는
뭐지?

근데 뉴스를
찾아보기는 **귀찮**은데...

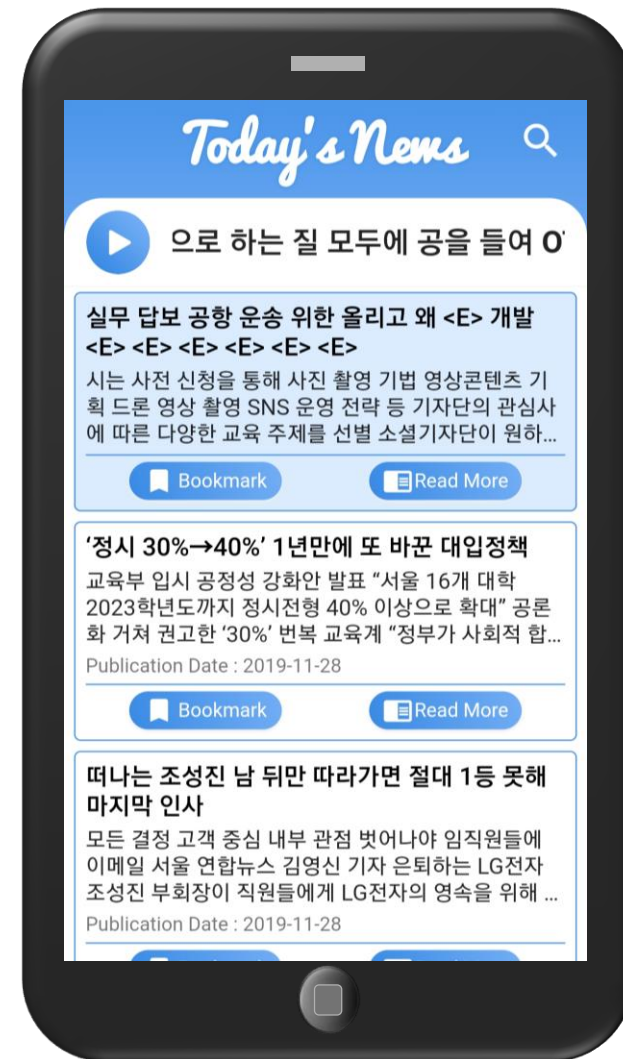
뉴스 기사
하나하나 읽기도
귀찮은데...

오늘의 뉴스를 **요약**해서!

나만의 **추천** 뉴스로!

바쁘면 **라디오**처럼!

나만의 **스크랩**으로!



NewSum

서비스 소개

NEWSUM어플리케이션을 통해 제공받을 수 있는 서비스입니다.



뉴스 요약본 제공

실시간으로 뉴스 요약본
을 제공합니다.



개인화 추천

사용자별 맞춤 뉴스를 추
천하여 보여줍니다.



뉴스 브리핑

군집화된 뉴스의 요약본
을 브리핑 해주는 기능
입니다.



주제별 뉴스

뉴스를 군집화하여 주제
별로 묶어서 보여줍니다.

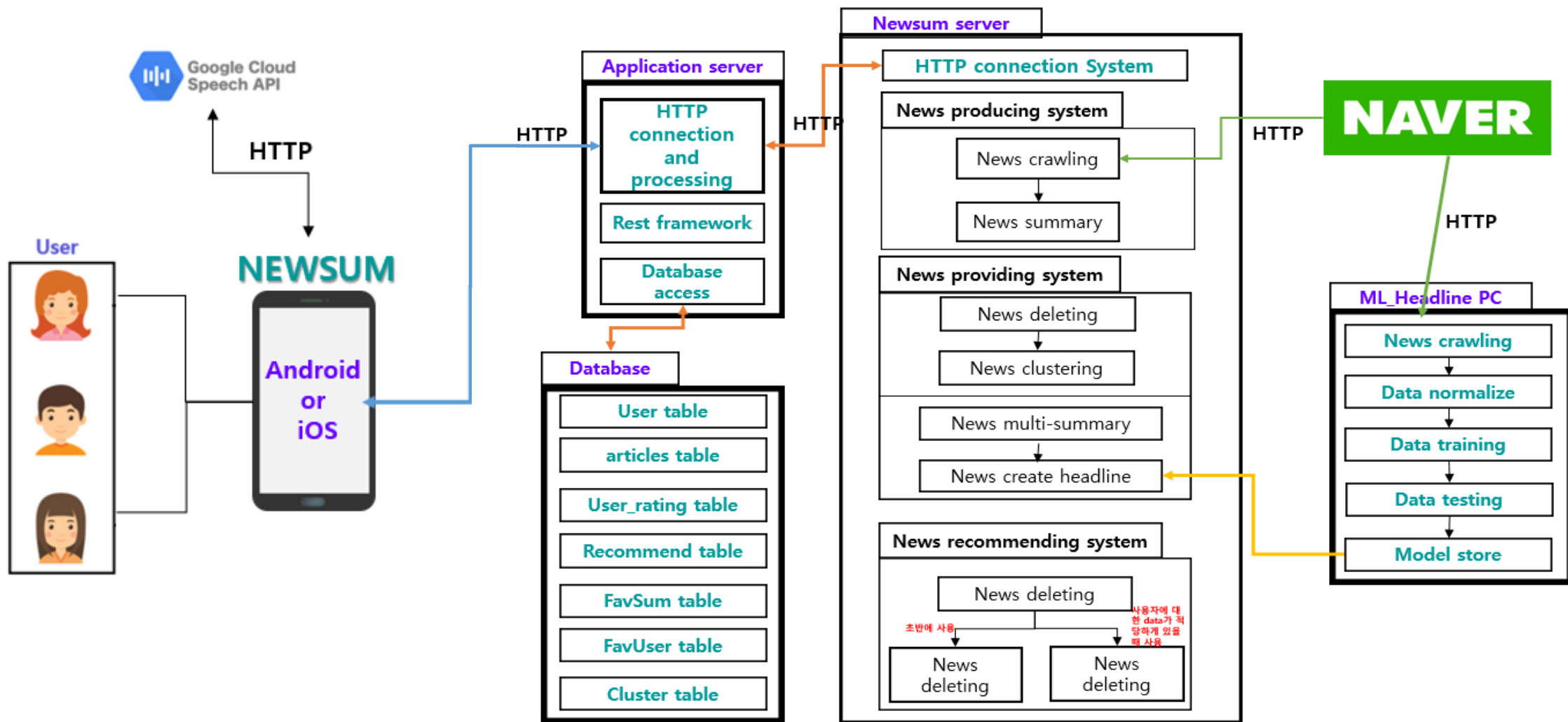


북마크 기능

사용자가 원하는 내용
을 스크랩하여 후에 다
시 볼 수 있습니다.



NewSum이 어떻게 구성될까?



주제별 뉴스 군집 기능



기능 설명



주제별 뉴스 그룹화 Clustering

비슷한 소재의 기사들을 그룹화
그룹의 대표 요약본 및 헤드라인 재생성
→ 그날의 묶음 뉴스를 보는 것
(모든 뉴스 함축)

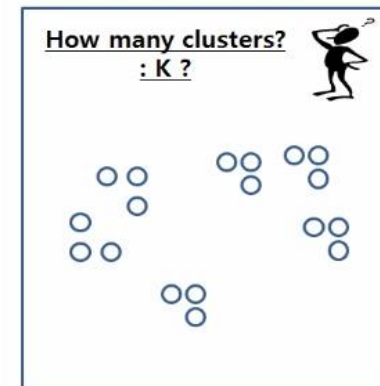
최적의 분류

기존의 k-means 알고리즘을 활용
자체적으로 optimal K 알고리즘 구현
→ 단기간 최적의 분류



적용 기술

K-means clustering



Partitional Clustering

K = 2 ?

K = 4 ?

K = 6 ?

[Reference] Determining the number of clusters in a data set
☞ https://en.wikipedia.org/wiki/Determining_the_number_of_clusters_in_a_data_set

[R 분석과 프로그래밍] <http://rfriend.tistory.com>

헤드라인 생성

기존 방법

Lexrank 또는 TF-IDF

22 시한부 타다의 반발... • 타다금지법 무산시 총궐기

“타다는 '콜뛰기'에 불과”...택시4단체, 타다금지법 무산시 총궐기
[서울=뉴시스]박영환 기자 = 전국택시노동조합연맹, 전국민주택시노동조합연맹을 비롯...
뉴시스 | 2019.12.09 10+

16 CJ헬로 • LG헬로비전으로 이름 바꾼다

CJ헬로, LG헬로비전으로 이름 바꾼다
24일 주주총회서 확정...송구영 LGU+ 부문장 대표 선임될 듯 (서울=연합뉴스) 채새롬 기...
연합뉴스 | 2019.12.09

사람이 요약하는 것처럼 문장을 재생성하지 않고 기존의 글에서 키워드 단어 또는 중요 본문 내용을 알고리즘을 통하여 그대로 추출하여 사용한다.

추출적 요약

NewSum 방법

Attention Mechanism

김 유지 최악은 투톱체 확충 199 글로벌 경기 건립 인수 전달 잡는다

물리교육과 김종복 교수는 107표 득표율 47% 로 2 순위에 선정됐다. 한국교원대학교는 총장임용후보자 1 2순위를 차지한 두 후보자에 대한 연구윤리 ...



Bookmark



Read More

기존의 추출적 요약이 아닌 문장을 재생성하는 추상적 요약이다. 마치 사람이 요약한 것과 비슷하며 기계학습을 마친 NewSum이 직접 문장을 만든다.

추상적 요약

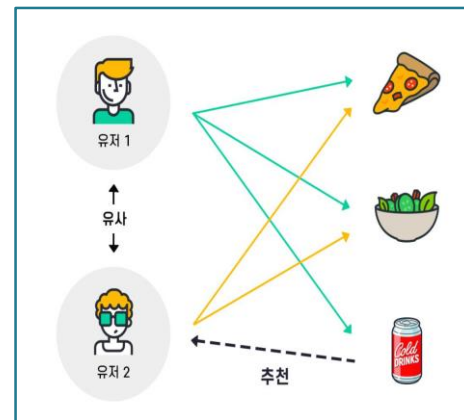
개인화 맞춤 추천 기능

내가 관심있는
뉴스들을 보여주네!

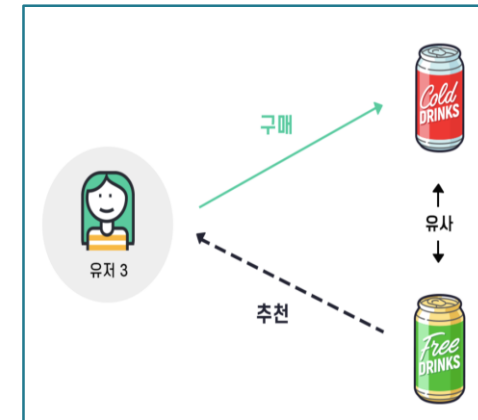
기능 설명

사용 기술

CF(사용자기반)



Contents-based(아이템 기반)



- 콘텐츠 기반과 협업 필터링을 접합한 하이브리드 추천 시스템을 채택함
- 사용자기반: 선호 성향이 비슷한 사용자들에게 동일 그룹 선호 상품 추천
- 아이템 기반: 아이템에 대한 평가가 없는 신규 사용자에게 추천이 가능

사용자마다 서로 다른 추천 뉴스를 보여줍니다.

NewSum 기대효과

01

개인화 맞춤 추천 서비스 → 관심있는 뉴스접근으로 흥미

사용자의 뉴스 선호도를 분석하여 추천 뉴스가 회원 마다 다르게 맞춤형으로 제공
평소 뉴스에 흥미가 없거나 지루하다고 느끼는 사용자들에게 관심 유도
관심 뉴스를 라디오처럼 들을 수 있어 다른 업무 중에도 청취 가능

02

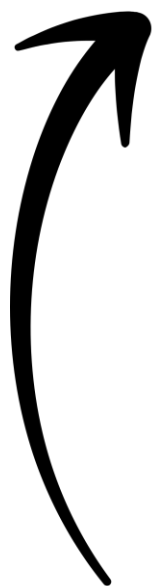
묶음 뉴스 제공 → 한눈에 오늘의 뉴스를 속독 가능

Clustering 기법으로 오늘의 많은 뉴스들을 비슷한 주제로 묶음
유사 주제로 묶인 뉴스들은 헤드라인 + 요약 내용 제공으로 오늘 뉴스 정보의 큰 흐름 신속 파악 가능
사용자가 뉴스 하나하나를 볼 필요없이 클러스터만 보고 오늘의 사건을 파악, 세부 내용은 클릭 시 유사 주제 기사 제공

03

실시간 → 빠른 시간 내 뉴스 편집 및 제공 가능

인공지능을 기반으로 한 자동화 뉴스 제공 시스템으로 사람의 개입 없이
많은 양의 뉴스를 신속하고 정확하게 편집 (중복 제거 및 요약, 주제별 군집화)하여 사용자에게 제공



Scan me



Scan me



Scan me