

RecoNews

Spring Boot + Python Side Project

RecoNews

1. 프로젝트 소개

1-1. 웹 프로젝트 소개

1-2. 개발환경

2. 시스템 설계

2-1. DB 테이블 정의서

2-2. 사용자 흐름 차트

3. 주요화면 기능

UI/UX 및 핵심 기능 설명

4. 결과 · 회고

1. 프로젝트 소개

1-1. 웹 프로젝트 소개



뉴스 크롤링 엔진과 Spring Boot 기반의 사용자 관리 시스템이 결합된 뉴스 서비스 플랫폼



개발 기간

2025.12 ~ 2026.02 (주 2-3회 진행)



기획 배경

뉴스 홍수 시대에 필요한 핵심 정보 요약 및 개인화 저장 서비스의 필요성



주요 기능

뉴스 자동 수집, 본문 요약,
회원별 관심 뉴스 저장 및 관리

1. 프로젝트 소개

1-2. 개발 환경

Python → Spring REST API → MySQL

Backend & API



Spring Boot

: RESTful API 설계 및 비즈니스 로직 구현

Spring Data JPA

: 객체 지향적 쿼리 및 MySQL 데이터베이스 연동

Data Processing



Python 3.x

: 뉴스 데이터 수집 및 전처리 파이프라인 구축

Newspaper3k & Feedparser

: 뉴스 본문 자동 추출 및 RSS 피드 파싱

Frontend



TypeScript / HTML5 / CSS3 / JavaScript

: 타입 안정성을 확보한 사용자 UI/UX 구현

Database



MySQL

: 사용자 정보 및 뉴스 메타데이터 저장

2. 시스템 설계

2-1. DB 테이블 정의서

• Users

회원 정보 관리 테이블

Col	Column Name	Data Type	Length	Key	Null 여부	ETC	Comments
1	id	int		PK	N	AUTO_INCREMENT	회원 계정 식별 번호
2	email	varchar	100	UK	N		회원 이메일
3	created_at	datetime			Y	CURRENT_TIMESTAMP	계정 생성일

 2. 시스템 설계

2-1. DB 테이블 정의서

• News

기사 정보 관리 테이블

Col	Column Name	Data Type	Length	Key	Null 여부	ETC	Comments
1	id	int		PK	N	AUTO_INCREMENT	기사 식별 번호
2	title	varchar	500		N		기사 제목
3	summary	text			N		기사 요약 내용
4	url	varchar	500	UK	N		기사 원문 주소
5	created_at	datetime		MUL	Y	CURRENT_TIMESTAMP	기사 수집 시각
6	image_url	varchar			Y		기사 이미지 주소

2. 시스템 설계

2-1. DB 테이블 정의서

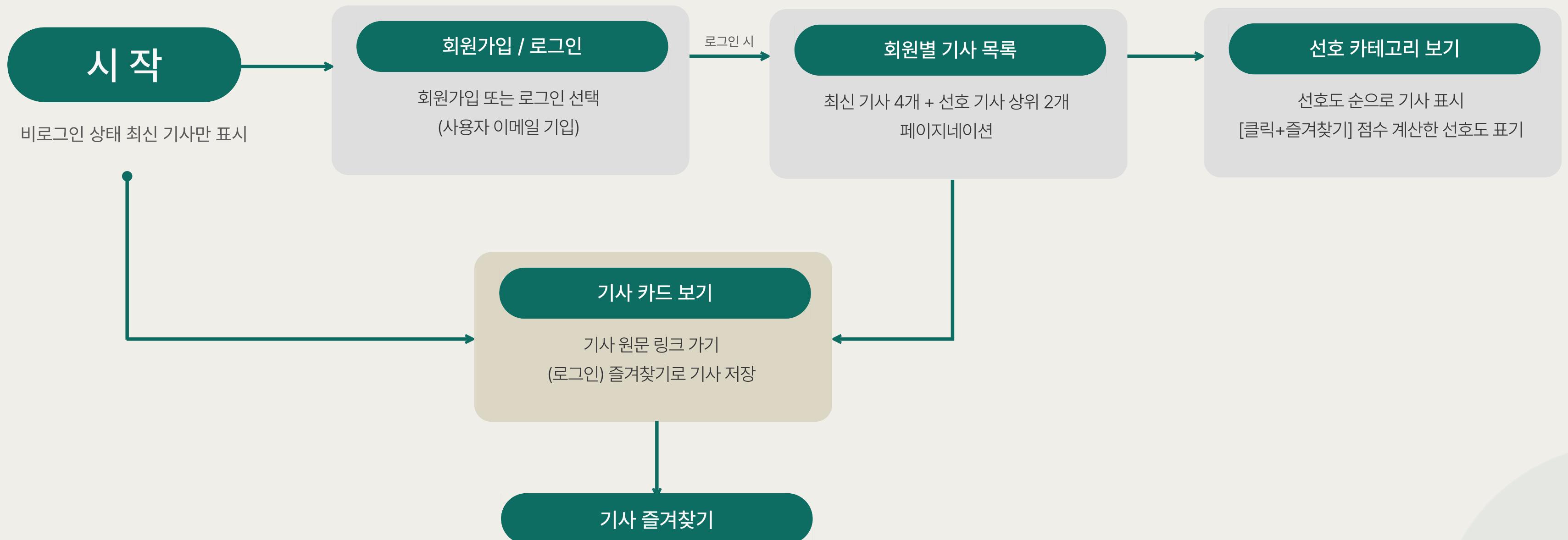
• User_actions

회원별 개인 선호도 기사 관리 테이블

Col	Column Name	Data Type	Length	Key	Null 여부	ETC	Comments
1	id	int		PK	N	AUTO_INCREMENT	식별 번호
2	user_id	int		MUL	N		회원 계정 식별 번호
3	news_id	int		MUL	N		기사 식별 번호
4	action_type	enum('CLICK','LIKE')			N		행동 유형
5	created_at	datetime			Y	CURRENT_TIMESTAMP	행동 일시

2. 시스템 설계

2-2. 사용자 흐름 차트



3. 주요 화면 기능

시작 화면 (비로그인 상태)

1. RecoNews

최신 뉴스

4

5

1 2 3 Signup Login

[속보] 구광모 회장, LG그룹 상속재산 분쟁 승소
구광모 LG그룹 회장을 상대로 어머니와 여동생 등 LG 오너일가가 제기한 상속회복청구 소송 1심에서 구 회장이 승소했습니다.

'AI 거풀론' 아랑곳없이...미국 억만장자 투자자, 메타에 신규 투자
빌 애크먼 [로이터 연합뉴스 자료사진] [로이터 연합뉴스 자료사진] 미국의 억만장자 투자자 빌 애크먼이 이끄는 헤지펀드 퍼싱 스퀘어가 '인공지능(AI) 거풀론'에도 불구하고 페이스북·스타그램 운영사인 메타에 상당한 규모의 지분을 투자한 것으로

"송민호, 복무 기간 430일 중 102일 이탈"...관리자도 가담 정황
재판매 및 DB 금지] [YG엔터테인먼트 제공. 그룹 위너의 멤버 송민호(32)가 사회복무요원 복무 당시 근무시간 중 복무 이탈한 일수기 총 102일에 달하는 것으로 파악됐습니다.

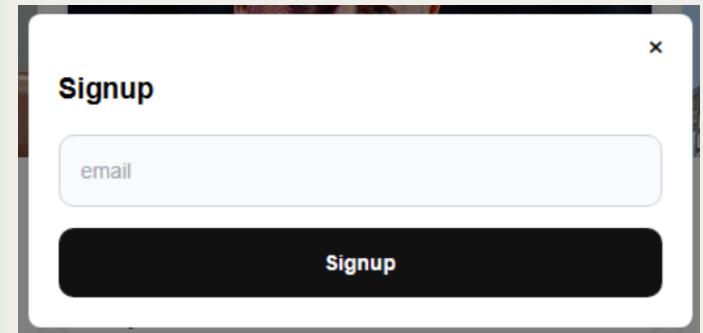
군, '코브라' 이어 '수리운' 헬기도 비행 중단..."긴급 점검"
재판매 및 DB 금지] [육군 제공. 군 당국이 국산 기동헬기 '수리운' 비행을 전면 중단하고 긴급 안전 점검을 진행 중인 것으로 알려졌습니다.

장동혁 "李대통령 오찬 참석 재논의"...국힘 최고위 공개 반대
이 대통령과의 오찬 재고 의사 밤힌 장동혁 대표 이 대통령과의 오찬 재고 의사 밤힌 장동혁 대표 (서울=연합뉴스) 황광모 기자 = 국민의힘 장동혁 대표가 12일 국회에서 열린 최고위원회...에서 이재명 대동령, 더불어민주당 정청래 대표와의 오찬에 불

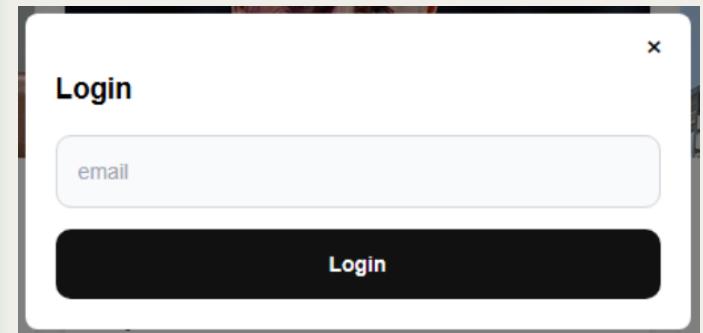
본 서비스에 표시되는 기사/이미지의 저작권은 각 언론사 및 콘텐츠 제공자에게 있으며, 본 서비스는 기사 링크 및 일부 정보(제목/요약 등)를 제공하는 목적으로입니다.
© 2026 news

1 2 3 4 5 6 7 8 9 10 >

2. 회원가입 모달



3. 로그인 모달



4. 기사 카드 (이미지 미제공 시)



No	UI	기능 설명
1	로고	클릭 시 1페이지 이동
2	회원가입 버튼	이메일 입력으로 간단 회원가입
3	로그인 버튼	이메일 입력으로 간단 로그인
4	기사 카드	기사 이미지 + 기사 제목 + 요약한 기사 내용 (이미지 미제공 시 대체 이미지 사용)
5	페이지	페이지네이션

3. 주요 화면 기능

로그인 화면

RecoNews

최신 뉴스

1 test@test.com 2 favorite 3 Logout 4 내가 자주 보는 기사

[속보] 구광모 회장, LG그룹 상속재산 분쟁 승소
구광모 LG그룹 회장 [연합뉴스 자료사진] [연합뉴스 자료사진] 구광모 LG그룹 회장을 상대로 어머니와 여동생 등 LG 오너일가가 제기한 상속회복청구 소송 1심에서 구 회장이 승소했습니다.

'AI 거품론' 아랑곳없이...미국 억만장자 투자자, 메타에 신규 투자
빌 애크먼 [로이터 연합뉴스 자료사진] [로이터 연합뉴스 자료사진] 미국의 억만장자 투자자 빌 애크먼이 이끄는 헤지펀드 퍼싱 스퀘어가 '인공지능(AI) 거품론'에도 불구하고 페이스북·인스타그램 운... 사이인 메타에 상당한 규모의 지분을 투자한 것으로 확인됐다고

쿠팡투자자 한국 정부 상대 소송에 미 투자회사 3곳 추가참여
쿠팡 본사 [연합뉴스 자료사진 제공] [연합뉴스 자료사진 제공] 쿠팡의 미국 투자자들이 쿠팡의 개인정보 유출 사건에 대한 한국 정부의 대응을 문제 삼아 추진하는 법적 다툼이 확대되고 있습니다.

"송민호, 복무 기간 430일 중 102일 이탈"...관리자도 가담 정황
재판매 및 DB 금지] [YG엔터테인먼트 제공. 그룹 위너의 멤버 송민호(32)가 사회복무요원 복무 당시 근무시간 중 복무 이탈한 일수가 총 102일에 달하는 것으로 파악했습니다.

연합뉴스 TV
지난달 러시아 공격으로 파괴된 키이우 건물 [로이터=연합뉴스 제공] [로이터=연합뉴스 제공] 도널드 트럼프 미국 대통령과 블로디미르 철렌스키 우크라이나 대통령의 종전안 담판을 하루 가량 앞두... 러시아가 우크라이나 수도 키이우에 대규모 공습을 가했습니다.

연합뉴스 TV
은행 벽에 뚫린 거대한 구멍 [로이터] 은행 벽에 뚫린 거대한 구멍 [로이터] 크리스마스 연휴 기간, 독일에서 도둑들이 한 시중은행 벽을 드릴로 뚫고 금고를 털어가는 초유의 절도 사건이 벌어졌습니다.

본 서비스에 표시되는 기사/이미지의 저작권은 각 언론사 및 콘텐츠 제공자에게 있으며, 본 서비스는 기사 링크 및 일부 정보(제목/요약 등)를 제공하는 목적입니다.
© 2026 news

1 2 3 4 5 6 7 8 9 10 »

No	UI	기능 설명
1	이메일 정보	회원 이메일 정보
2	즐겨찾기 버튼	클릭 시 [내가 자주 보는 기사] 페이지로 이동
3	로그아웃 버튼	클릭 시 로그아웃
4	내가 자주 보는 기사	선호도 점수가 가장 높은 상위 2개 기사 표시

3. 주요 화면 기능

[내가 자주 보는 기사] 화면

RecoNews

test@test.com favorite Logout

내가 자주 보는 기사

 해당 기사는 이미지를 제공하지 않습니다.

연합뉴스 TV

지난달 러시아 공격으로 파괴된 키이우 건물 [로이터=연합뉴스 제공] [로이터=연합뉴스 제공] 도널드 트럼프 미국 대통령과 볼로디미르 짤렌스키 우크라이나 대통령의 종전안 담판을 하... 기량 앞두고 러시아가 우크라이나 수도 키이우에 대규모 공습

선호도: 22



연합뉴스 TV

은행 벽에 뚫린 거대한 구멍 [로이터] 은행 벽에 뚫린 거대한 구멍 [로이터] 크리스마스 연휴 기간, 독일에서 도둑들이 한 시중 은행 벽을 드릴로 뚫고 금고를 털어가는 초유의 절도 사건이 ... 어졌습니다.

선호도: 17



【속보】 구광모 회장, LG그룹 상속재산 분쟁 승소

구광모 LG그룹 회장 [연합뉴스 자료사진] [연합뉴스 자료사진] 구광모 LG그룹 회장을 상대로 어머니와 여동생 등 LG 오너일 가가 제기한 상속회복청구 소송 1심에서 구 회장이 승소했습니다.

선호도: 8

연합뉴스 TV

지난해 12월 한국프로야구선수협회 시상식인 '2025컴투스프로야구 리얼 글러브 어워드'에서 포수상을 받은 NC 김형준의 모습. [사진 제공] 연합뉴스 지난해 12월 한국프로야구선수회 시상식인 '2025컴투스프로야구 리얼 글러브 어워드'에서 포

선호도: 5



연합뉴스 TV

독일 한 자동차 판매업체의 SNS 화면 [서경덕 교수 SNS 캡처. 재판매 및 DB 금지] [서경덕 교수 SNS 캡처.] 평창 사륜 오토바이 주탁사고 [강원 특별자치도소방본부 제공]

선호도: 5



본 서비스에 표시되는 기사/이미지의 저작권은 각 언론사 및 콘텐츠 제공자에게 있으며, 본 서비스는 기사 링크 및 일부 정보(제목/요약 등)를 제공하는 목적입니다.

© 2026 news

1 2 3 4 5 6

```
@Query(
    value = """
        SELECT
            n.id,
            n.title,
            n.summary,
            n.image_Url,
            COALESCE(
                (
                    COALESCE(SUM(
                        CASE
                            WHEN ua.action_type = 'CLICK' THEN 1
                            WHEN ua.action_type = 'LIKE' THEN 3
                            ELSE 0
                        END
                    ), 0)
                +
                CASE
                    WHEN n.created_at >= NOW() - INTERVAL 1 DAY THEN 2
                    ELSE 0
                END
            ), 0
        ) AS total_score
    FROM news n
    LEFT JOIN user_actions ua
    ON ua.user_id = :userId
    AND ua.news_id = n.id
    GROUP BY n.id
    HAVING total_score > 0
    ORDER BY total_score DESC, n.created_at DESC
    LIMIT :limit
    """
    nativeQuery = true
)
```

선호도 점수 계산

- 1) 기사 클릭 시 +1, 즐겨찾기 버튼 클릭시 +3
- 2) 24시간 이내 기사일 경우 +2

해당 행동을 계산해서 선호도 점수 표기 및 내림차순으로
기사 목록 출력

3. 주요 화면 기능

기사 화면

RecoNews test@test.com favorite Logout

1 쿠팡투자자 한국 정부 상대 소송에 미 투자회사 3곳 추가참여

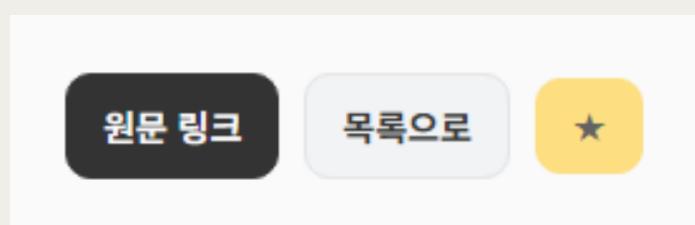


쿠팡 본사 [연합뉴스 자료사진 제공] [연합뉴스 자료사진 제공] 쿠팡의 미국 투자자들이 쿠팡의 개인정보 유출 사건에 대한 한국 정부의 대응을 문제 삼아 추진하는 법적 다툼이 확대되고 있습니다.

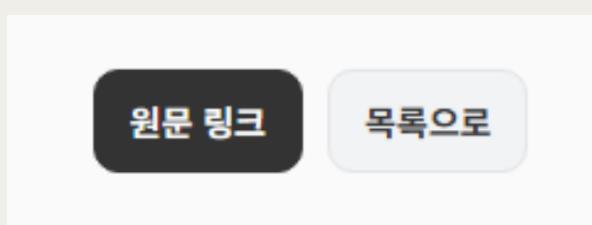
2 원문 링크 3 목록으로 ★

본 서비스에 표시되는 기사/이미지의 저작권은 각 언론사 및 콘텐츠 제공자에게 있으며, 본 서비스는 기사 링크 및 일부 정보(제목/요약 등)를 제공하는 목적입니다.
© 2026 news

· 즐겨찾기 클릭 시



· 비로그인 시



· 기사 요약 로직

```
# 사진/캡션 제거
def clean_text(text: str):
    text = re.sub(r'^.*?\n', '', text)
    text = re.sub(r'\(사진=.*?\)', '', text)
    return text.strip()

# 한국어 뉴스 문장 분리
def split_korean_sentences(text: str):
    # 1단계: 줄바꿈 기준으로 문단 분리
    paragraphs = re.split(r'\n+', text)

    sentences = []
    for para in paragraphs:
        para = para.strip()
        if len(para) < 20:
            continue

    # 2단계: 마침표 기준 문장 분리
    parts = re.split(r'(?<=\. )', para)
    for p in parts:
        p = p.strip()
        if len(p) >= 20:
            sentences.append(p)

    return sentences
```

```
# 영상 기사 스kip
def is_invalid_article(text: str) -> bool:
    if not text:
        return True

    if "video 태그를 지원하지 않습니다" in text:
        return True

    if len(text.strip()) < 100:
        return True

    return False
```

No	UI	기능 설명
1	기사 요약 정보	회원 이메일 정보
2	원문 링크 버튼	클릭 시 기사 원문으로 이동
3	즐겨찾기 버튼	1) 즐겨찾기 전: 빈 별(☆) 2) 즐겨찾기 후: 채워진 별(★), 버튼 클릭 비활성화 3) 비로그인 시: 해당 버튼 비활성화

기사 요약 로직

뉴스 URL → 본문 수집 → 불필요한 텍스트 제거 → 문장 분리 → 앞부분 요약 생성 → 영상/짧은 기사 필터링

- 1) 뉴스 원문 크롤링
newspaper3k 라이브러리 사용: 제목 / 본문 텍스트 추출
- 2) 기사 대표 이미지 추출: 뉴스 썸네일용 이미지 추출
- 3) 기사 제목 추출
- 4) 텍스트 정리 (사진/캡션 제거): 자주 등장하는 불필요 정보 제거
- 5) 한국어 문장 분리: 길이 20자 이상 문장만 사용
- 6) 요약 길이 제한
- 7) 영상 기사 스kip: 영상 전용 기사 or 본문 없는 기사 필터링

4. 결과 · 회고

● 결과

주요 성과 3가지

1	2	3
개인화 뉴스 추천 엔진 구현	뉴스 요약 파이프라인 자동화	사용자 중심 흐름 설계
내용 클릭 및 좋아요 데이터 기반 가중치 알고리즘 적용	내용 기사 수집 → 전처리 → 핵심 2문장 요약 프로세스 구축	내용 비로그인과 로그인 유저를 구분한 맞 춤형 서비스 동선 구축
성과 사용자 행동 로그를 점수화하여 최신성 (24시간 가산점)과 선호도가 결합된 맞 춤형 피드 제공	성과 영상 기사 등 요약이 불필요한 콘텐츠 자 동 필터링을 통해 뉴스 소비 효율성 극대 화	성과 로그인 시 최신 4개 + 선호 2개 구성의 개인화 메인 화면 제공

● 회고

아쉬운 점 및 향후 계획 3가지

1	2	3
요약 알고리즘의 한계 및 개선	통합 검색 기능 도입 (미구현)	검색어 기반 추천 알고리즘 확장
아쉬운 점 단순 앞문장 추출 방식 사용 시 기사 맥락에 따라 핵심 내용이 누락되는 문제 발생	아쉬운 점 초기 기획에는 포함되었으나, 일정 및 우선순위로 인해 검색 기능을 구 현하지 못함	아쉬운 점 현재는 클릭/좋아요 기반이며, 사용 자가 직접 검색한 키워드 정보는 활 용하지 못함
추후 계획 문장 길이 필터링을 넘어 BERT 기반 요 약 모델로 품질 고도화	추후 계획 기사 제목 및 본문을 포괄하는 전체 키 워드 검색 엔진 추가 개발	추후 계획 사용자가 자주 검색한 단어를 로그화하 여 추천 가중치에 반영하는 모델로 고도 화

" 사용자 행동을 읽는 추천 기술, 더 깊이 있는 정보 전달을 위한 첫걸음을 뗈습니다.
 기술적 보완과 기능 확장을 통해, 누구에게나 '가장 필요한 뉴스'만 전하는 플랫폼으로 나아가겠습니다. "