

# 푸른거북 X 플라이북

20191567 김유진

20191688 한창훈

20183416 홍석진

# 목차

BOOKA 소개(시연 영상)

별점 기반 추천 알고리즘(Collaborative Filtering)

컨텐츠 기반 추천 알고리즘(Contents-based Filtering)

기대효과



**BOOKA 시연**

# 필요성

문화·라이프 > 책

## 책 더 안 읽는다...성인 10명 중 4명, 1년 독서량 '0'

뉴스스

입력 2020.03.11 13:52



## [취재K] 영화관람 세계1위, 독서는 OECD 꼴찌...‘문화기형’ 숨은 이유

입력 2019.02.15 (18:39)

취재K

4 29



책을 읽고는 싶는데,  
어떤 책을 읽지?



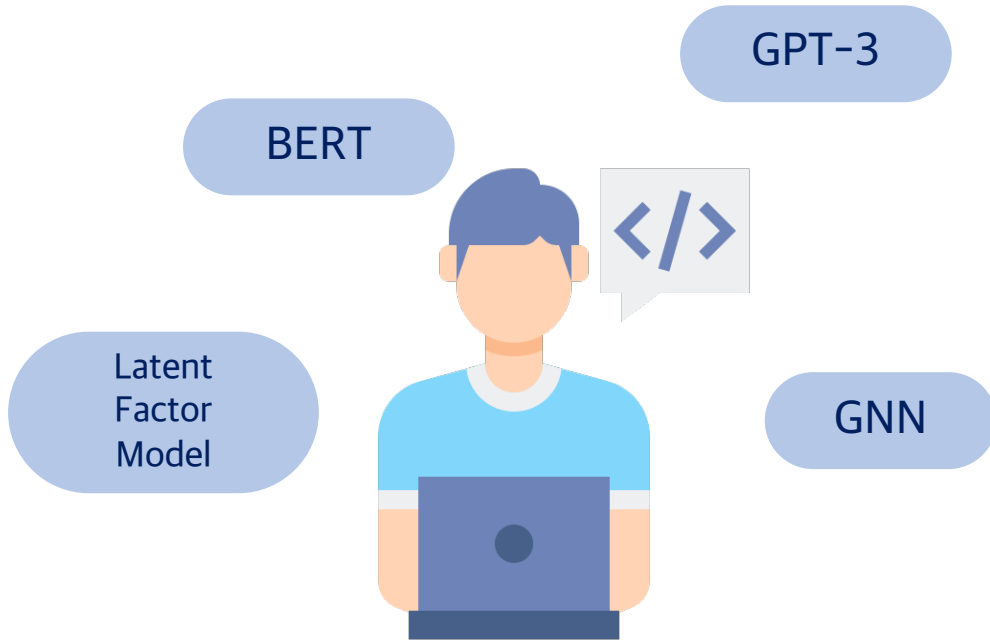
▶ 책을 **고르다가** 흥미를 잃음

▶ **적절하지 못한** 책을 읽고 흥미를 잃음

<https://news.kbs.co.kr/news/view.do?ncd=4139652>

[https://www.chosun.com/site/data/html\\_dir/2020/03/11/2020031102172.html](https://www.chosun.com/site/data/html_dir/2020/03/11/2020031102172.html)

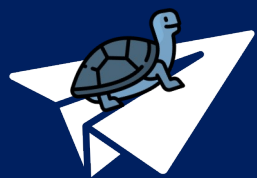
# 목표



플라이북의 추천 시스템을 발전시키기 위해  
다양한 접근 시도

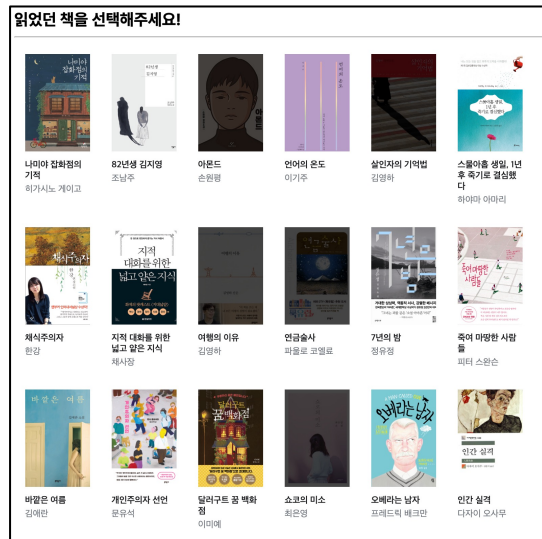


도서 추천 서비스를 만들어 보며  
추천 시스템을 적용



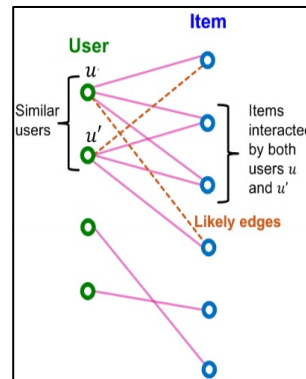
# 별점 기반 추천 알고리즘

# 별점 기반 추천 알고리즘 - LightGCN

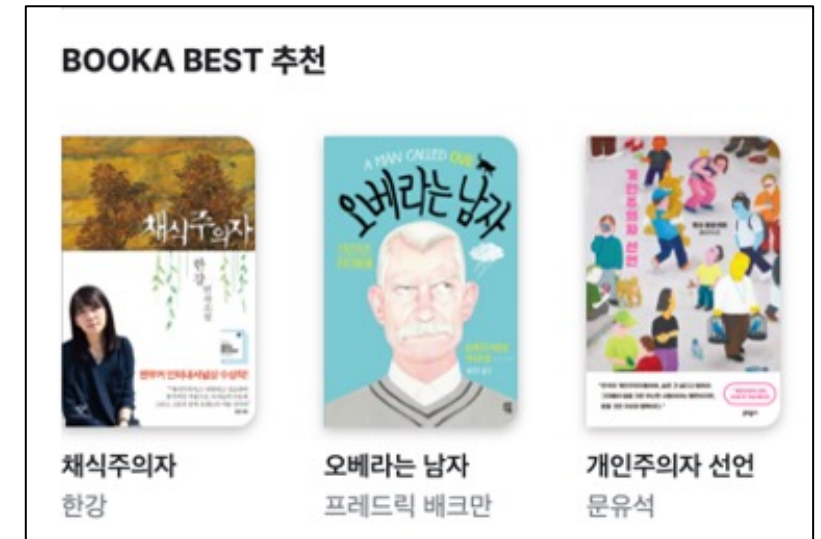


첫 로그인 시에 사용자가  
읽었던 책을 선택

## LightGCN



LightGCN을 이용해  
유저의 취향 파악



읽은 책을 기반으로 추천한다.

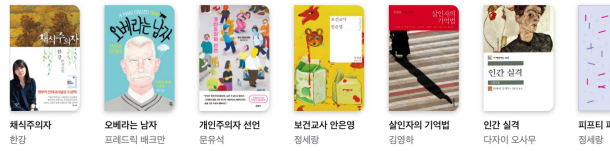
# 별점 기반 추천 알고리즘 - LightGCN

## Collaborative filtering

### 별점 기반 추천

LightGCN의 예측 별점이 높은 책 추천

#### BOOKA BEST 추천



#### 비슷한 사람이 읽은 책



#### 소년이 온다: 이 책과 비슷한 책



#### 당신만을 위한 추천 도서



### 나와 비슷한 취향을 가진 유저들이 읽은 책 추천

나와 유사도가 가장 높은 사용자 3명이 읽은 책 추천

### 읽은 책과 비슷한 책 추천

내가 읽은 책 중 하나를 뽑아 읽은 책과 유사도가 높은 책 추천

### 개인화된 추천

Global baseline을 적용해 이 유저의 특징을 크게 반영하여 더 개인화한 추천

유사도 계산은 모두 LightGCN의 결과인 임베딩을 이용하여 계산



# Latent Factor Model vs LightGCN

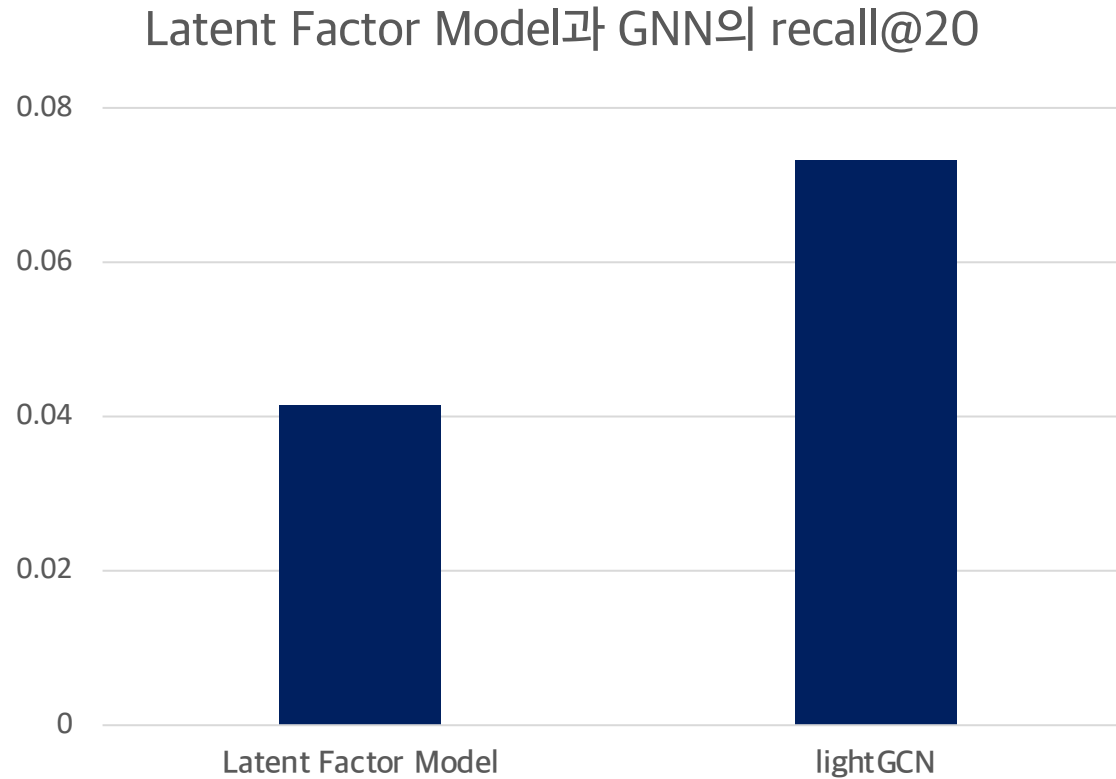
Latent Factor Model

**VS**

LightGCN



# Latent Factor Model vs LightGCN



LightGCN의 Recall이 Latent Factor Model 모델 보다 약 1.5배가 높음

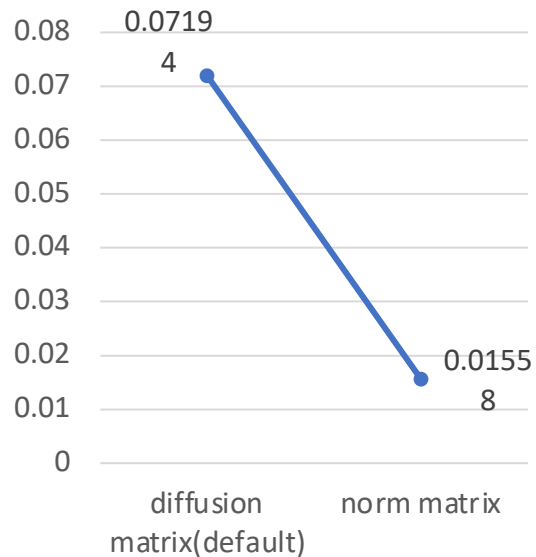


LightGCN 선택

# LightGCN 최적화

1

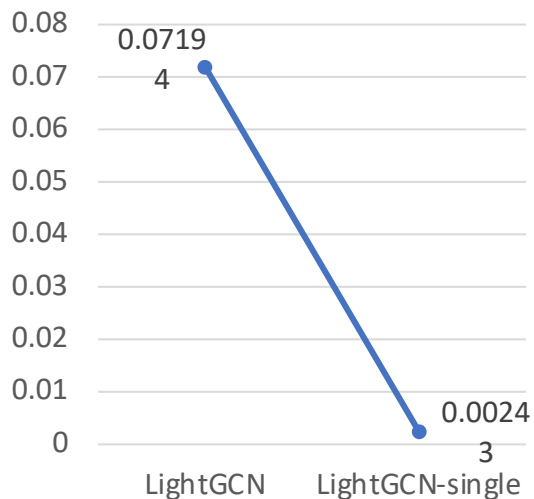
Matrix 종류에 따른 recall@20



Message passing 시  
normalization matrix 사용

2

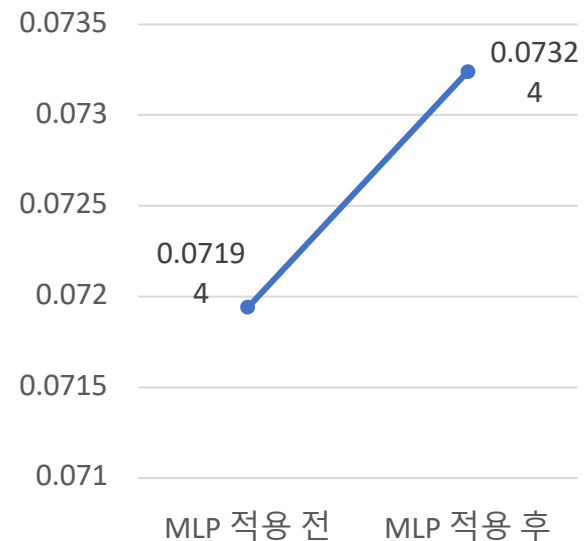
LightGCN과 LightGCN-single의  
recall@20 비교



LightGCN-single 사용

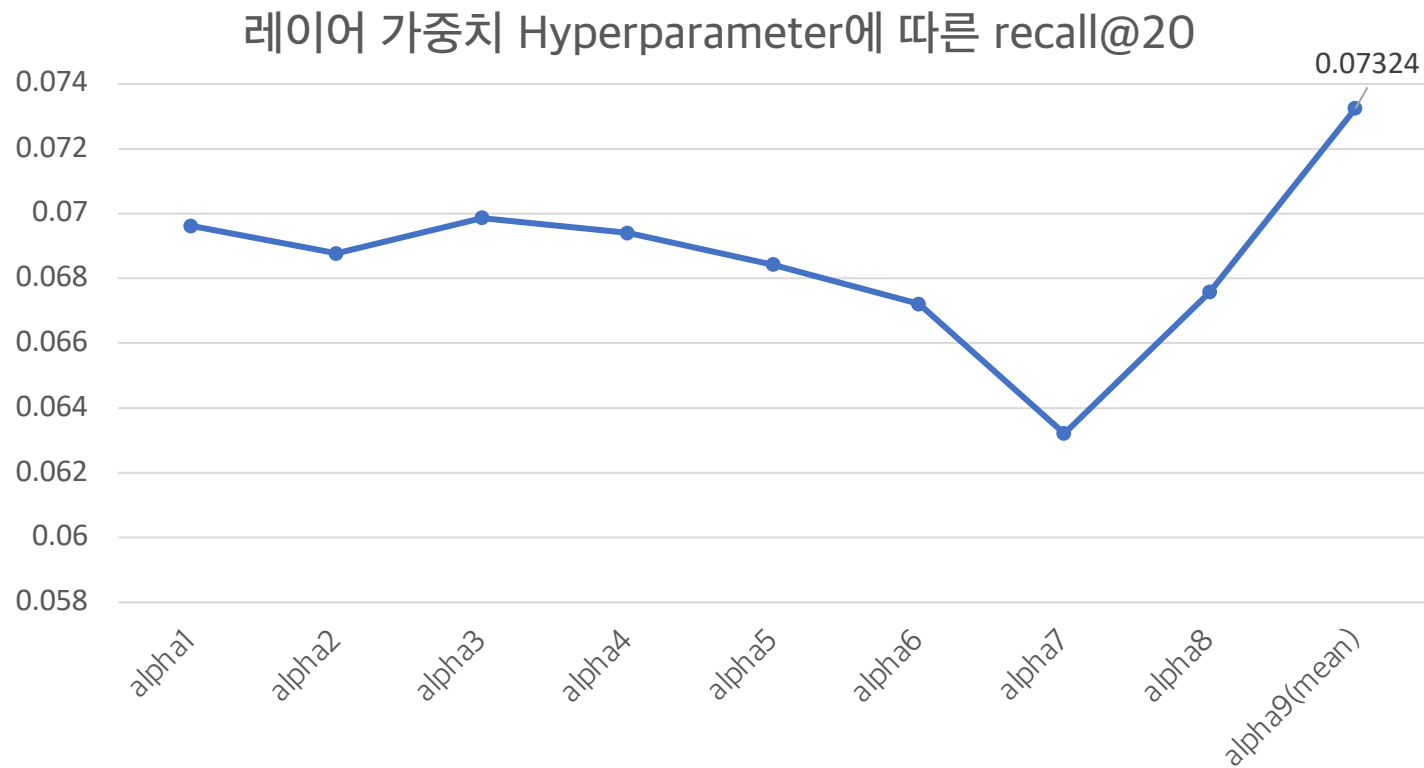
3

MLP 적용여부에 따른 recall@20



LightGCN 레이어 입력 전,  
MLP로 node feature를 전처리

# LightGCN 최적화



over-smoothing을 방지하기 위해 레이어 가중치를 조절



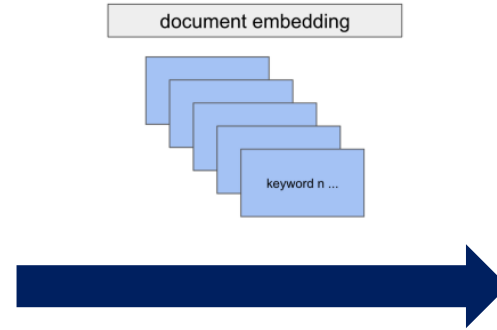
# 컨텐츠 기반 추천 알고리즘

# 컨텐츠 기반 추천 알고리즘 - 키워드 추출



책 소개
2004년 발표된 이래 수많은 독자들의 뜨거운 지지와 입소문으로 변함없이 맞아 새롭게 출간되었다. "처음 읽었을 때와 지금의 감상이 다른, 언제까지 사를 받
책 소개
2004년 발표된 이래 수많은 독자들의 뜨거운 지지와 입소문으로 변함없이 맞아 새롭게 출간되었다. "처음 읽었을 때와 지금의 감상이 다른, 언제까지 사를 받
책 소개
2004년 발표된 이래 수많은 독자들의 뜨거운 지지와 입소문으로 변함없이 맞아 새롭게 출간되었다. "처음 읽었을 때와 지금의 감상이 다른, 언제까지 사를 받으며 수많은 명대사 명장면을 탄생시킨 이 이야기는, 연애소설의 공
책 소개
30대 초중반, 어느덧 익숙해진 평범한 일상 속에서 사랑에 대한 설렘을 마 함 110호의 우편물>에는 우리와 크게 다르지 않은, 적당히 외로워하며 살 전히 많은 그들에게 사랑은 흔해빠진 감정이고 때로는 부질없어서 환멸이 그럼에도 불구하고 '다시 한 번 사랑해보기로 한' 그들의 이야기를 담담하 에 없는 우리를 위로해주고 있다. 독자들에게 의해 <사서함 110호의 우편물> 사하니까"로 끊임없이 회자되는 이 작품은, 10년이 넘는 시간이 지났음에도 고 있다.

## KeyBERT



KeyBert를 이용하여  
키워드 추출



### 1Q84 1

무라카미 하루키 지음 문학동네 펴냄

#순위

#마찬가지

#갑자기

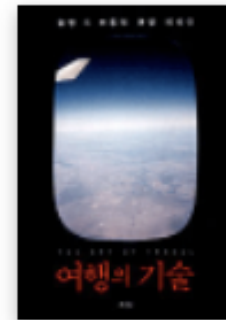
#추구

#훈련

#문예지

#먼저

#비밀



### 여행의 기술

알랭 드 보통 지음 이레 펴냄

#워즈워스

#안내

#예술

#시나이

#마드리드

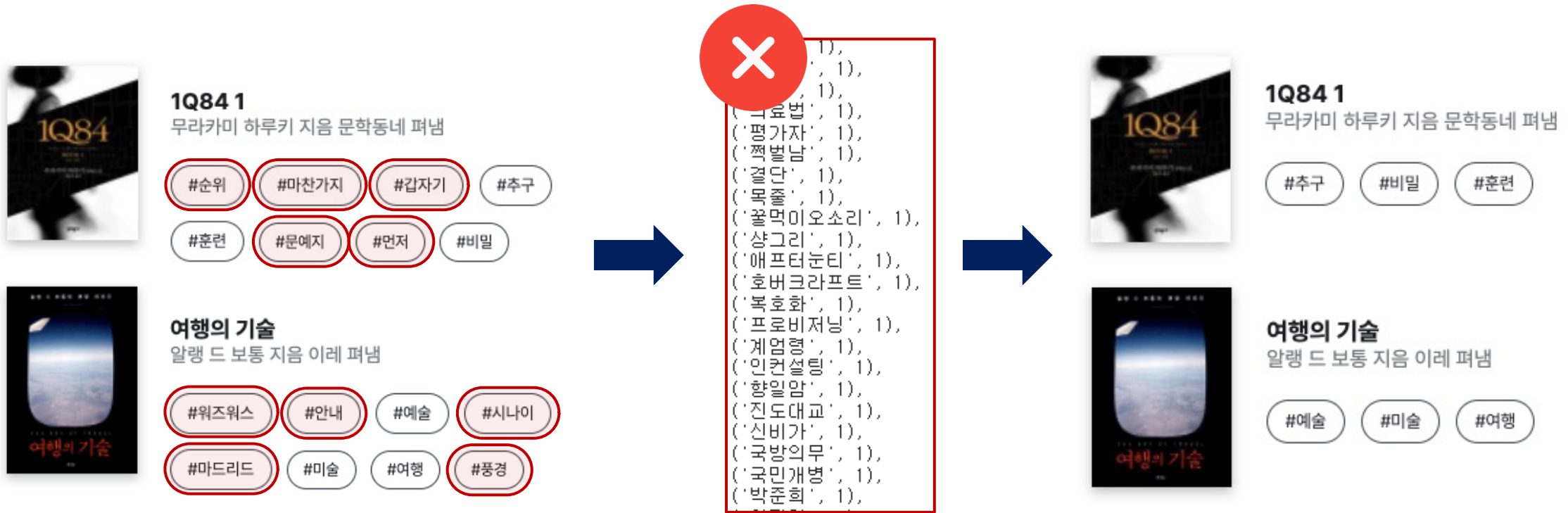
#미술

#여행

#풍경



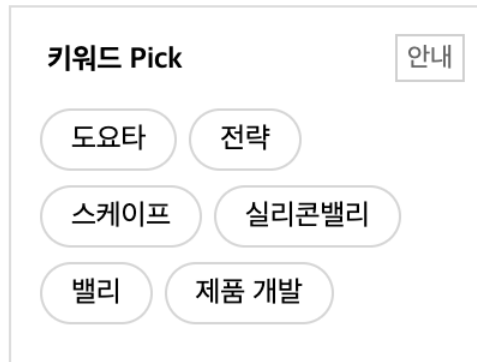
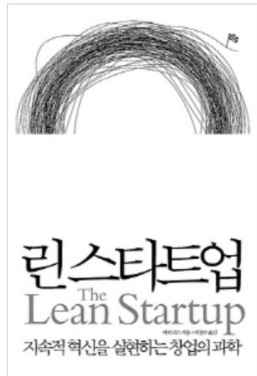
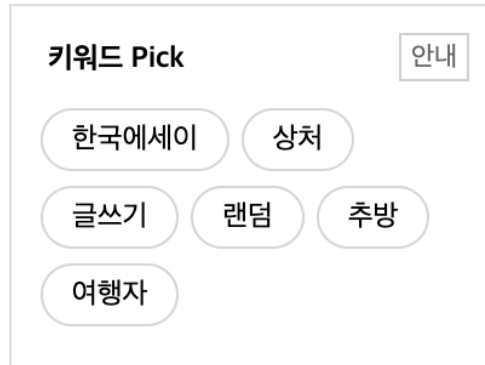
# 컨텐츠 기반 추천 알고리즘 - 키워드 정제



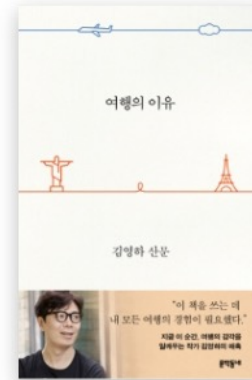
- ✓ 30번 이하로 등장한 키워드 제거
- ✓ 적절치 않은 키워드, 고유명사 등은 직접 정제
- ✓ 단어의 의미가 비슷한 키워드를 묶음 (ex 한국, 대한민국)



# 컨텐츠 기반 추천 알고리즘 - 키워드 결과 비교



교보문고의 키워드 픽



**여행의 이유**  
김영하 지음 문학동네 펴냄



**린 스타트업**  
에릭 리스 지음 인사이트 펴냄



**boka.**



# 컨텐츠 기반 추천 알고리즘

## Contents-based recommendation

### 상상 속 세상 #판타지



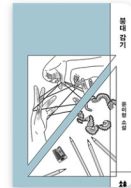
전이, 지니  
정유정



할머니가  
미안하다고  
전해달랬어요  
프레드릭 베르크만



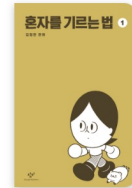
지금 이 순간  
기욤 뮈소



봉대 감기  
윤이형



상냥한 저승사자를  
기르는 법  
치넨 미키토



혼자서 기르는 법 1  
김정연



지금이  
초능력  
장강명

### 나를 돌보는 법을 잊어버린 당신에게



당신이 옳다  
정혜신



센서티브  
일자 샌드



나도 아직 나를  
모른다  
차미미



가짜 감정  
김용태



너의 목소리가  
들려  
김수현

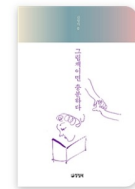


정확한 사랑의  
실험  
김수현



죽고 싶  
땀복아  
싶어 2  
김수현

### #동화책



그림책이면  
충분하다  
김영미



레온체로 도서관  
탈출게임  
크리스 그라번스...



미술가게  
너새니얼 호손...



움베르토 에코의  
지구를 위한 세  
가지 이야기  
움베르토 에코



어린 왕자  
앙투안 드 생텍...



컬트동화  
남희영



엄마 파  
있나요  
이영란

## 흥미로운 키워드를 가진 책 추천


책의 키워드 중

직접 흥미로운 키워드를 선별하여 추천

## 내가 읽은 책과 키워드를 공유하는 도서

사용자가 읽을 가능성이 큰 순서대로 정렬하여 표시

# 키워드 활용



**나미야 잡화점의 기적**

히가시노 게이고 지음 | 현대문학 펴냄


#소설
#기적
#사랑
#향기
#기도

♡ 읽고싶어요
✓ 읽었어요

**책 소개**

2012년 3월 일본에서 출간되어 큰 화제를 불러일으킨, 히가시노 게이고 최신작. 이번 이야기에는 그동안 히가시노 게이고 하면 떠올랐던 살인 사건이나 명탐정 캐릭터는 전혀 등장하지 않는다. 그럼에도 불구하고 마치 퍼즐을 맞추어가는 듯한 치밀한 짜임새는 과연 히가시노 게이고의 작품답게 명불허전의 짜릿한 쾌감을 선사하며 감동을 자아내 작가의 고정 독자를 충분히...


### '#알고리즘' 검색 결과



**인공지능의 현재와 미래**

나는 미래다 방송제작팀 지음 보아스 펴냄

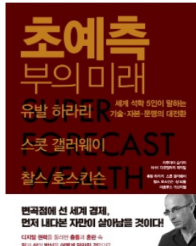
#알고리즘
#기술
#창업
#혁신



**인공지능 비즈니스 트렌드**

테크니들 지음 와이즈맵 펴냄

#알고리즘
#실리콘밸리
#혁신
#비즈니스
#스타트업



**초예측 부의 미래**

장 티롤 외 5명 지음 웅진지식하우스 펴냄

#경제학
#알고리즘
#기업인
#경제
#자본주의

키워드만 보고 유추할 수 있도록 책 정보에 노출

같은 키워드의 책 검색



# 기대효과

# 기대효과



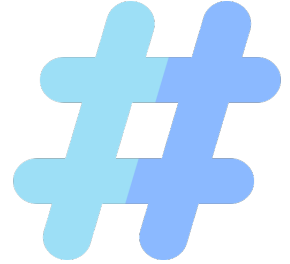
독서인구 증가



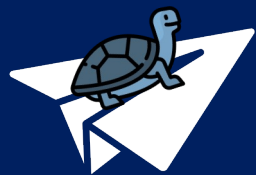
독서 편식 해결



도서 검색 개선

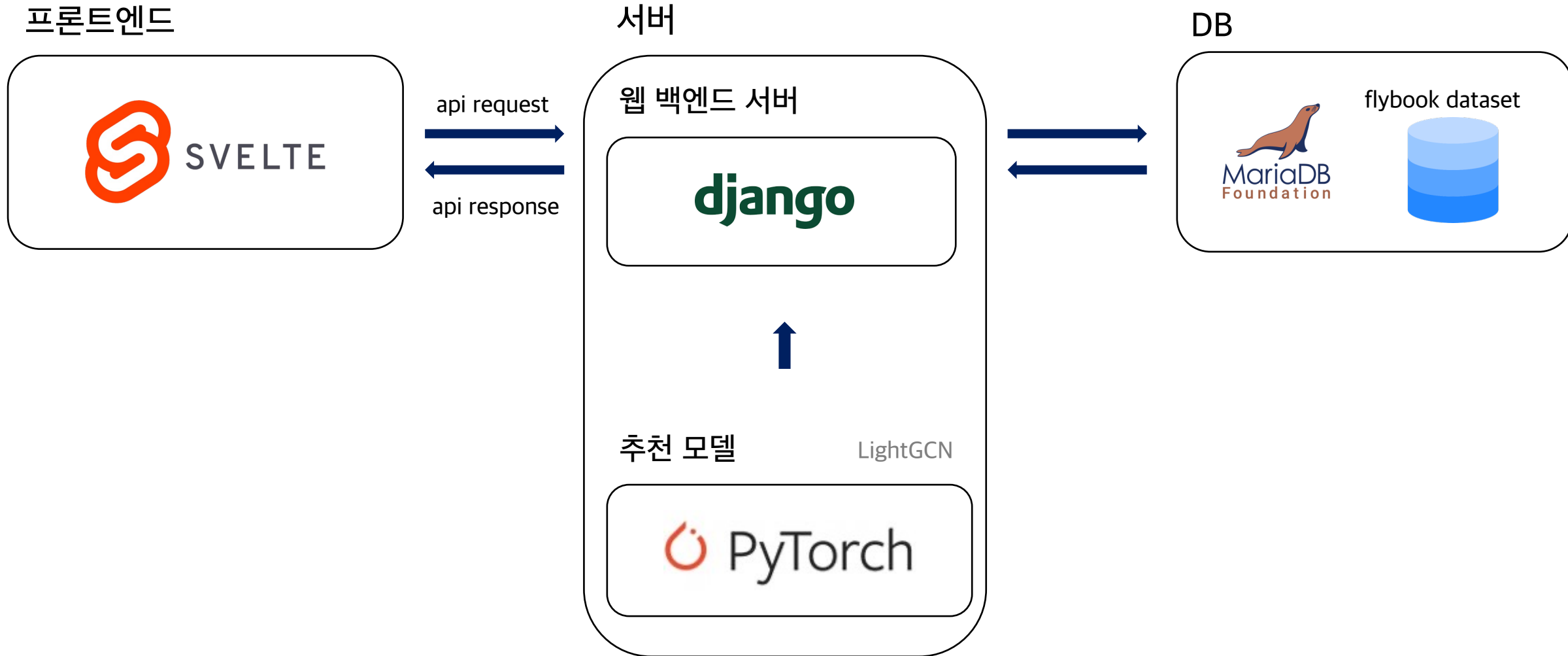


자동 태그 기능



# 기술 스택

# 시스템 구조



# 역할 분배

Frontend



김유진



**BOOKA 웹**  
책 추천 웹 서비스

Backend



한창훈



**BOOKA 서버**  
책 추천 시스템 / 웹 서버



**BOOKA DB**  
DB 설계 및 관리, 데이터 정리

AI Model



홍석진



**책 추천 시스템**  
책 추천 시스템 모델 개발



**키워드 추출**  
키워드 추출 시스템 모델 개발



공통

**책 추천 시스템**  
책 추천 시스템 설계

**키워드 추출**  
키워드 추출 시스템 설계

**서비스 기획**  
아이디어 / UI 디자인



감사합니다