

사랑의 홈쇼핑 창성핑

트렌드 기반 홈쇼핑 상품 추천 서비스

이유송(팀장), 장은별, 최창성, 송우선, 김수현, 정수빈

목차

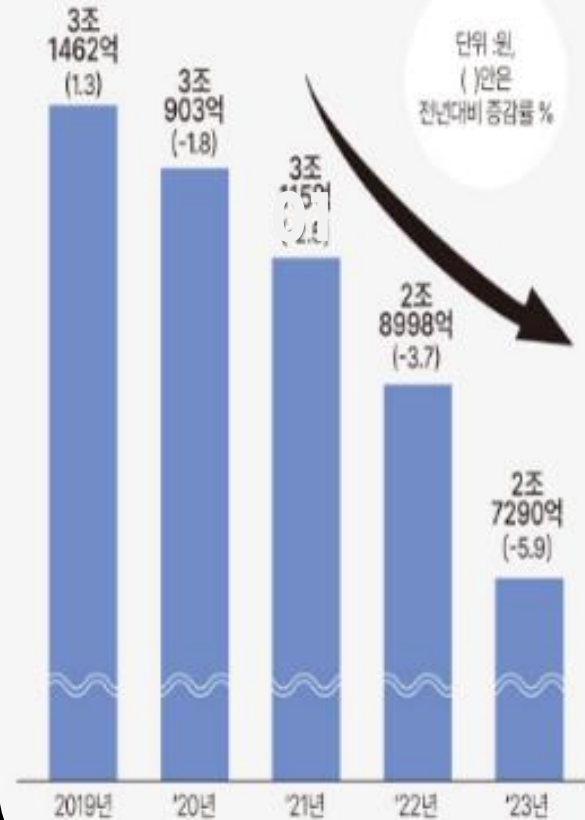


1. 프로젝트 개요

1. 배경
2. 목적 및 기대효과
3. 일정
4. 기술 스택

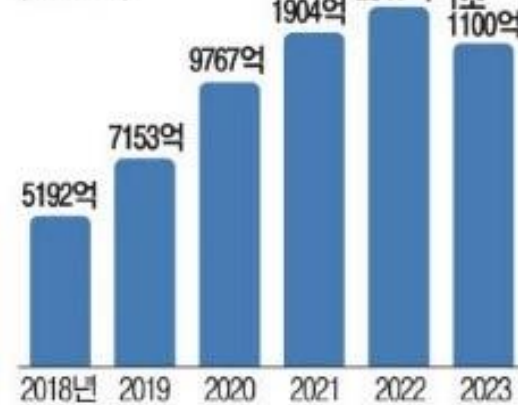
홈쇼핑 시장 동향

최근 5년간 TV홈쇼핑 7개사 방송매출액



DealSite

라이브 커머스사 매출 (단위=원)



*2023년은 추정치, SK스토아·KT알파·신세계라이브쇼핑·쇼핑엔터·W쇼핑 5개사 합산.

홈쇼핑

라이브 커머스의 발달로 지속적으로 감소하고 있는 홈쇼핑 시장.

커머스

계속해서 증가하고 있는 라이브 커머스 홈쇼핑의 시청률을 빼앗고 있는 중이다.

라이브 커머스 시장 규모



출처: 글로벌 리서치센터 / (주)비브이브리 '2023 라이브커머스 중점산업 리포트'

1. 배경

MZ 세대들의 소비행태

L (Loss of mass market)

소비의 양극화로 사라지고 있는 대중 시장

I (Immersive experience)

소유보다는 경험을 중시하는 소비자

Q (Question for efficiency)

돈 보다는 시성비/효율성이 중요한 소비

D (Digital innovation)

소비 환경의 패러다임 변화를 주도하는
디지털 혁신 기술

I (Intensifying wellness & sustainability)

핵심 키워드로 부상하는 웰니스와 지속 가능성

U (Unique value)

대중적인 흐름보다 개인의 차별화된 가치를 추구

리퀴드 소비

리퀴드(Liquid) 소비란?

소비 패턴이 고정되어 있지 않고 액체처럼 유동적으로
움직이기 때문에 예측 불가능한 소비

2. 목적 및 기대효과

트렌드 기반 홈쇼핑 상품 추천 서비스



구매 결정 용이

사용자가 선호할만한 제품을 먼저 제안하며 관련 상품 추천으로 추가 구매를 유도한다, 쇼핑 경험을 풍부하게 만든다.



최신 트렌드 반영

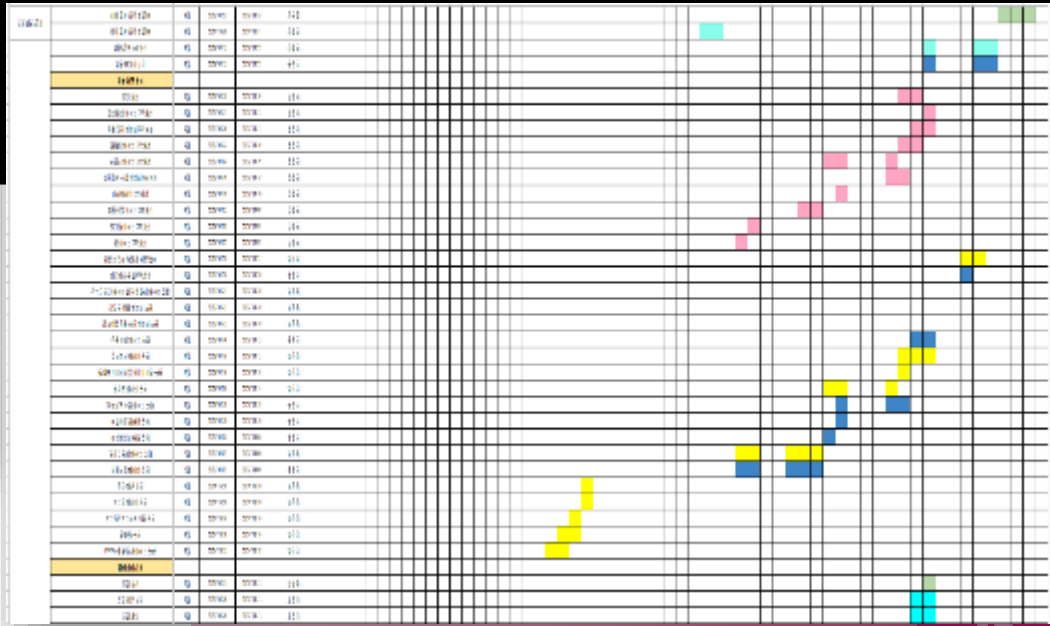
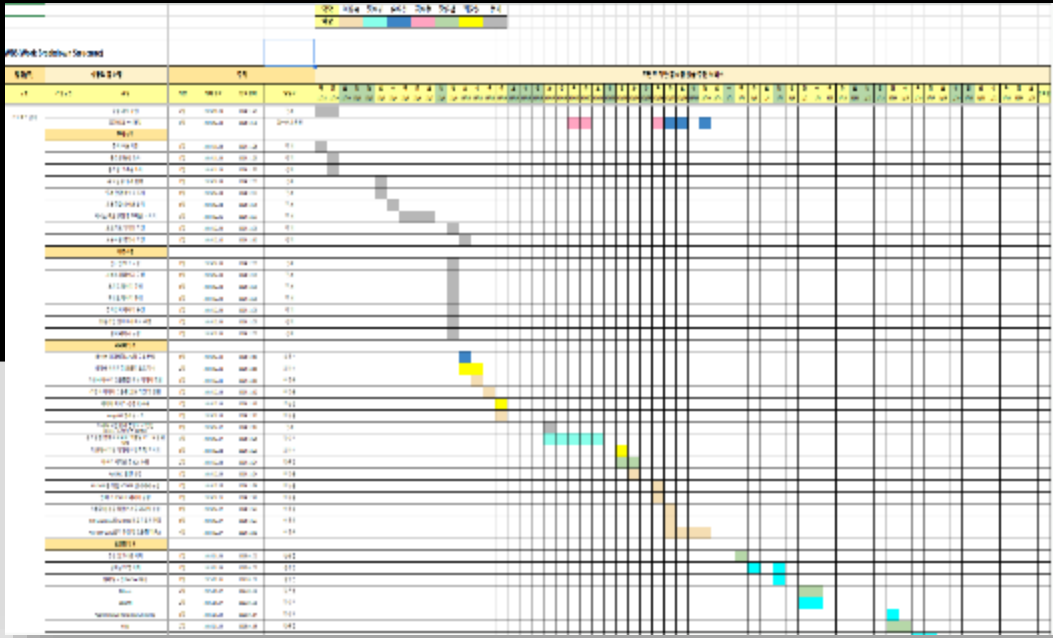
현재 유행하는 상품이나 인기 있는 제품을 추천 받아 트렌드 반영. 소비자들이 최신 유행 상품을 쉽게 접할 수 있다.



사용자 만족도 향상

사용자 맞춤형 서비스를 통해 개인화된 경험을 제공. 만족스러운 쇼핑 경험으로 재방문을 상승과 충성의 증가를 기대할 수 있다.

3. 일정




프로젝트 시작
2024-12-11


설계
2024-12-11
~2024-12-31


개발
2025-01-01
~2025-01-24


최종 발표
2025-01-25

4. 기술 스택

Front End



Back End



Infrastructure



Data Pipeline



DataBase



API

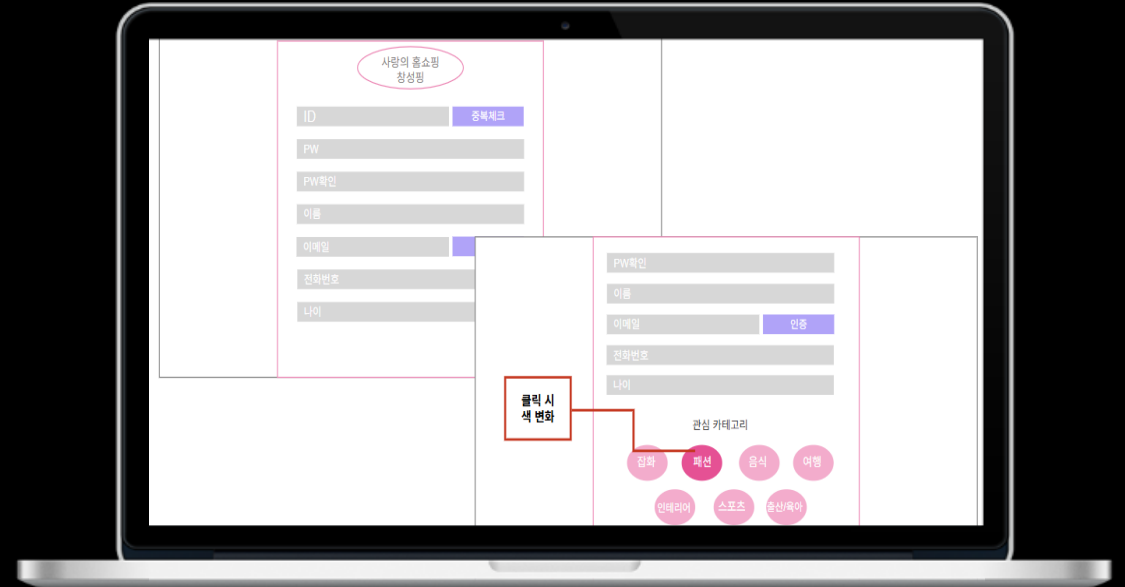
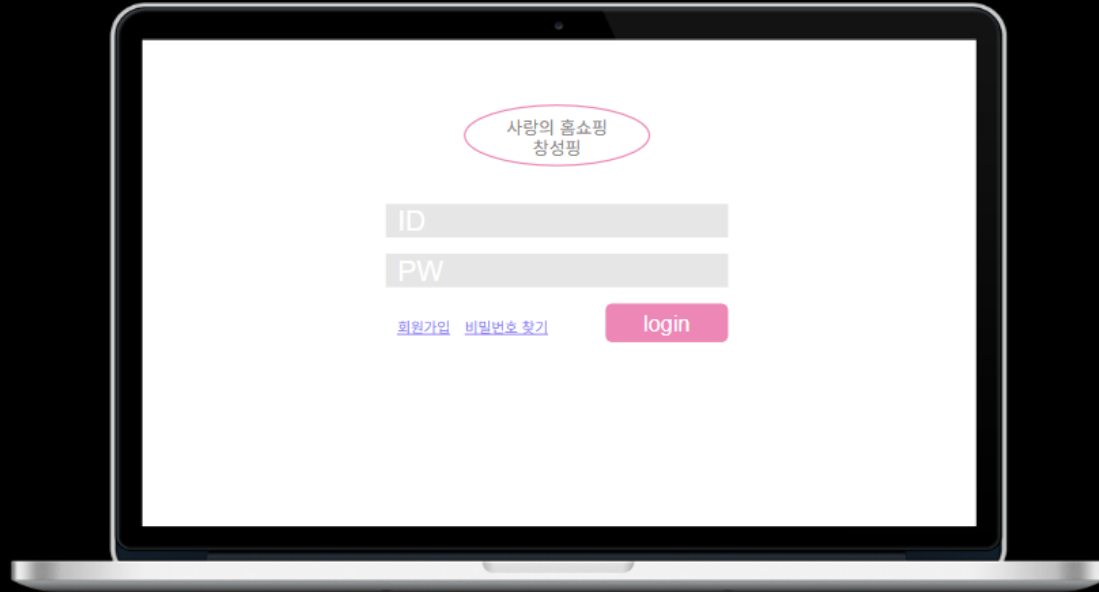


WIKIPEDIA
The Free Encyclopedia

2. 서비스 기획

1. 서비스 기획

1. 서비스 기획



1. 서비스 기획



3. 데이터 수집 및 전처리

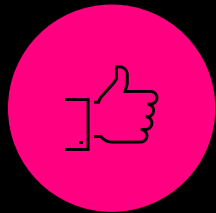
1. 수집할 데이터 선정
2. 데이터 파이프라인 구축 및 스케줄 생성
3. 전처리
4. 데이터 베이스 설계

1. 수집할 데이터 선정

네이버 트렌드 데이터

네이버데이터랩 쇼핑 인사이트 데이터

원본 : <https://datalab.naver.com/>



카테고리

패션의류

패션잡화

화장품/미용

디지털/가전

가구/인테리어

출산/육아

식품

스포츠/레저

생활/건강

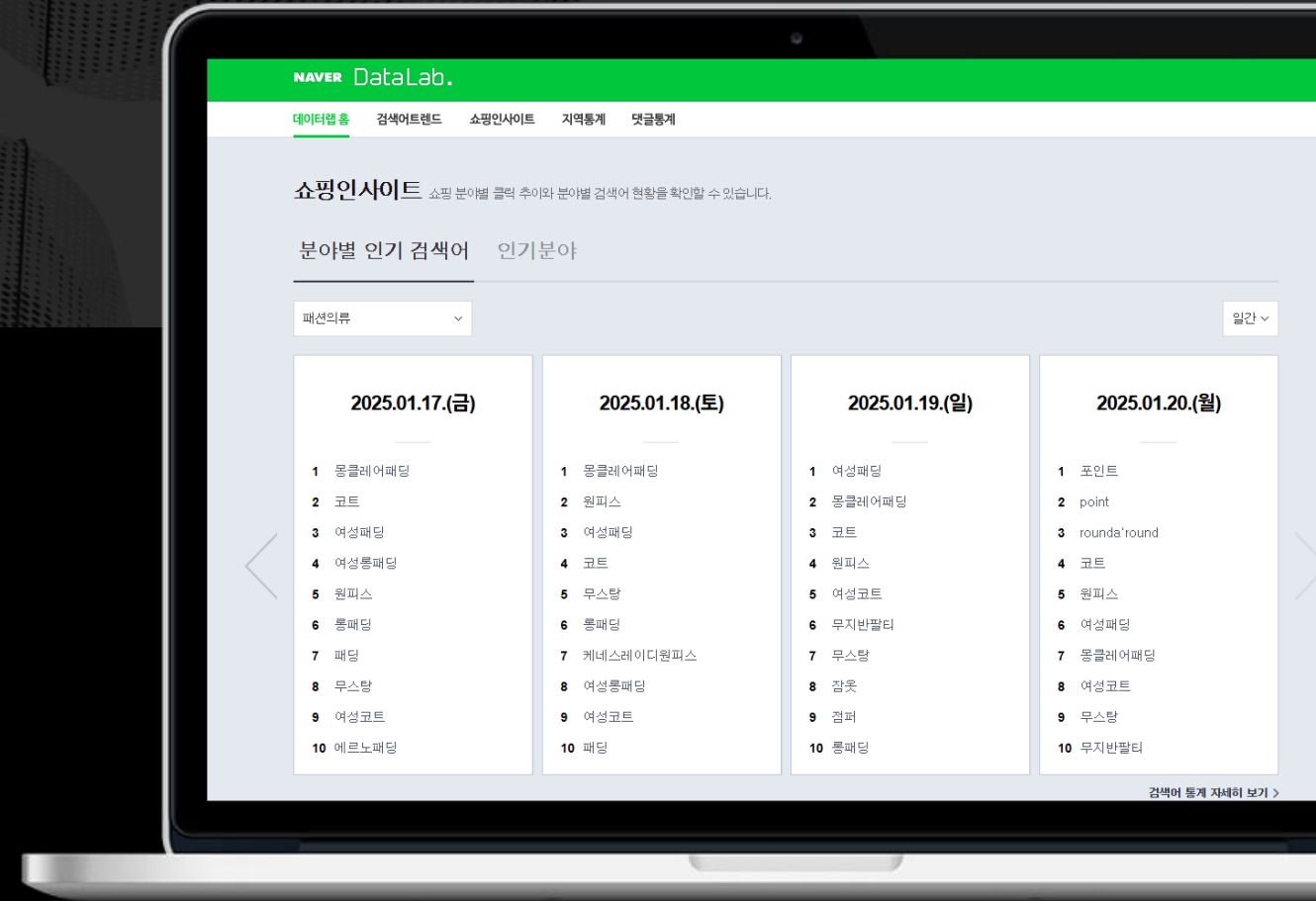
여가/생활편의



형태 및 기타

형태: 일간(12일), 주간(12주), 월간(12월)

데이터 건수: 7200건

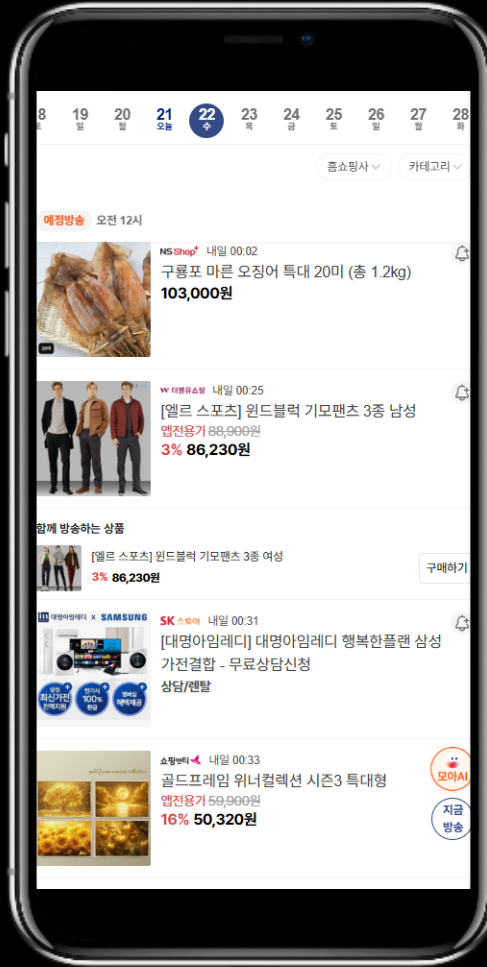


크롤링 코드

완성율 : 80%

여부 쇼핑물 편성 스케줄 일자 편성 완료 일수

<input type="checkbox"/>	NS 플러스	7일	3일
<input type="checkbox"/>	NS TV	7일	3일
<input checked="" type="checkbox"/>	KT 알파쇼핑	2일	
<input checked="" type="checkbox"/>	W 홈쇼핑	3일	
<input checked="" type="checkbox"/>	공영 홈쇼핑	10일	3일
<input checked="" type="checkbox"/>	GS Shop Live	10일	3일
<input checked="" type="checkbox"/>	GS myShop	10일	3일
<input checked="" type="checkbox"/>	홈&쇼핑	10일	4일
<input checked="" type="checkbox"/>	현대 TV	7일	2일
<input checked="" type="checkbox"/>	현대 플러스	7일	2일
<input checked="" type="checkbox"/>	롯데 TV	10일	2일
<input checked="" type="checkbox"/>	롯데 L-live	10일	2일
<input type="checkbox"/>	쇼핑엔티	4일	3일
<input checked="" type="checkbox"/>	SK 스토어	10일	3일
<input type="checkbox"/>	신세계 홈쇼핑	3일	3일
<input checked="" type="checkbox"/>	CJ TV	7일	3일
<input checked="" type="checkbox"/>	CJ PLUS	7일	2일



1. 수집할 데이터 선정

홈쇼핑 상품 데이터

데이터 원본: 각 홈쇼핑 사 편성표 (총 17개의 홈쇼핑사)

형태: 날짜, 방송시간, 방송날짜, 제목, 가격, 상품 URL, 이미지 URL (특정 홈쇼핑에 한해서 리뷰 수, 판매량, 별점 수집)

수집 기간: 홈쇼핑 별 상이(옆의 표 참조)

데이터 건수: 37500건(1월 22일 기준)

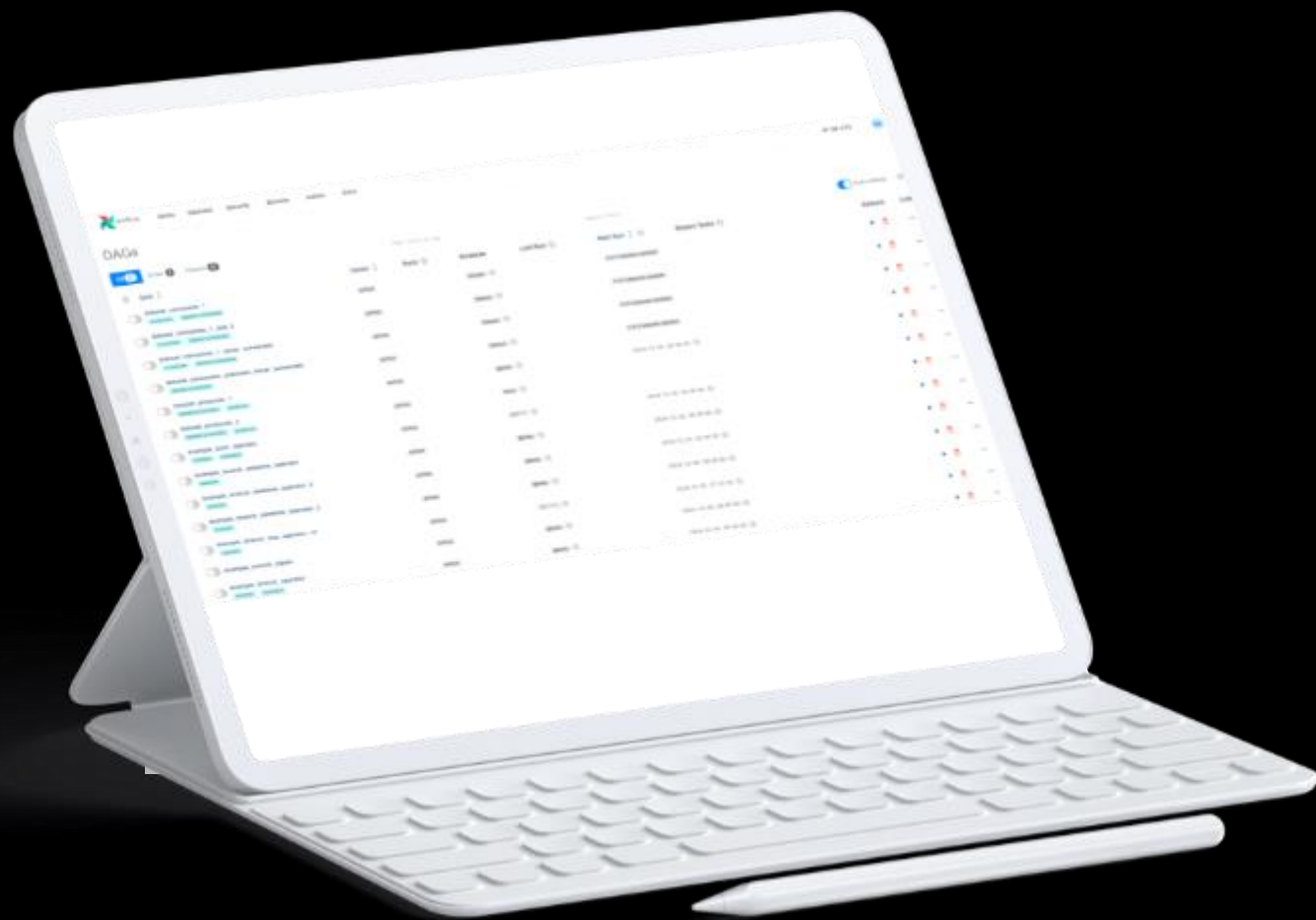


데이터 파이프라인 구축 AIRFLOW를 이용한 크롤링 코드 자동화

수집 주기: 1일

형식: 각 홈쇼핑과 트렌드 데이터

크롤링 Task를 넣어두고 자동화 실시



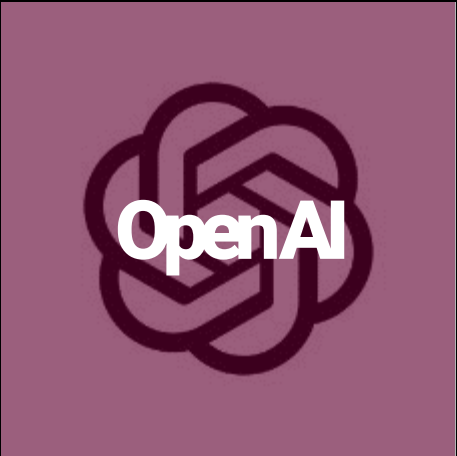
카테고리 키워드 추출 : Open AI

	title	Category
	조지루시 가열식 가슴기 특집 ✨	디지털/가전
온기플러스 조끼 엘라이트 특가 🔥 +5% 추가 할인 +10% 적립까지!		패션의류
[제나벨] 연말결산 특집 ❤️ 선착순 99특가 -70% SALE 파격혜택 📺		패션의류
베이비페어 ❤️ [블루래빗] 크리스마스 특집 🎄 ~68% SALE 파격혜택!		출산/육아
	베이비페어 ❤️ [알집매트]	출산/육아

카테고리

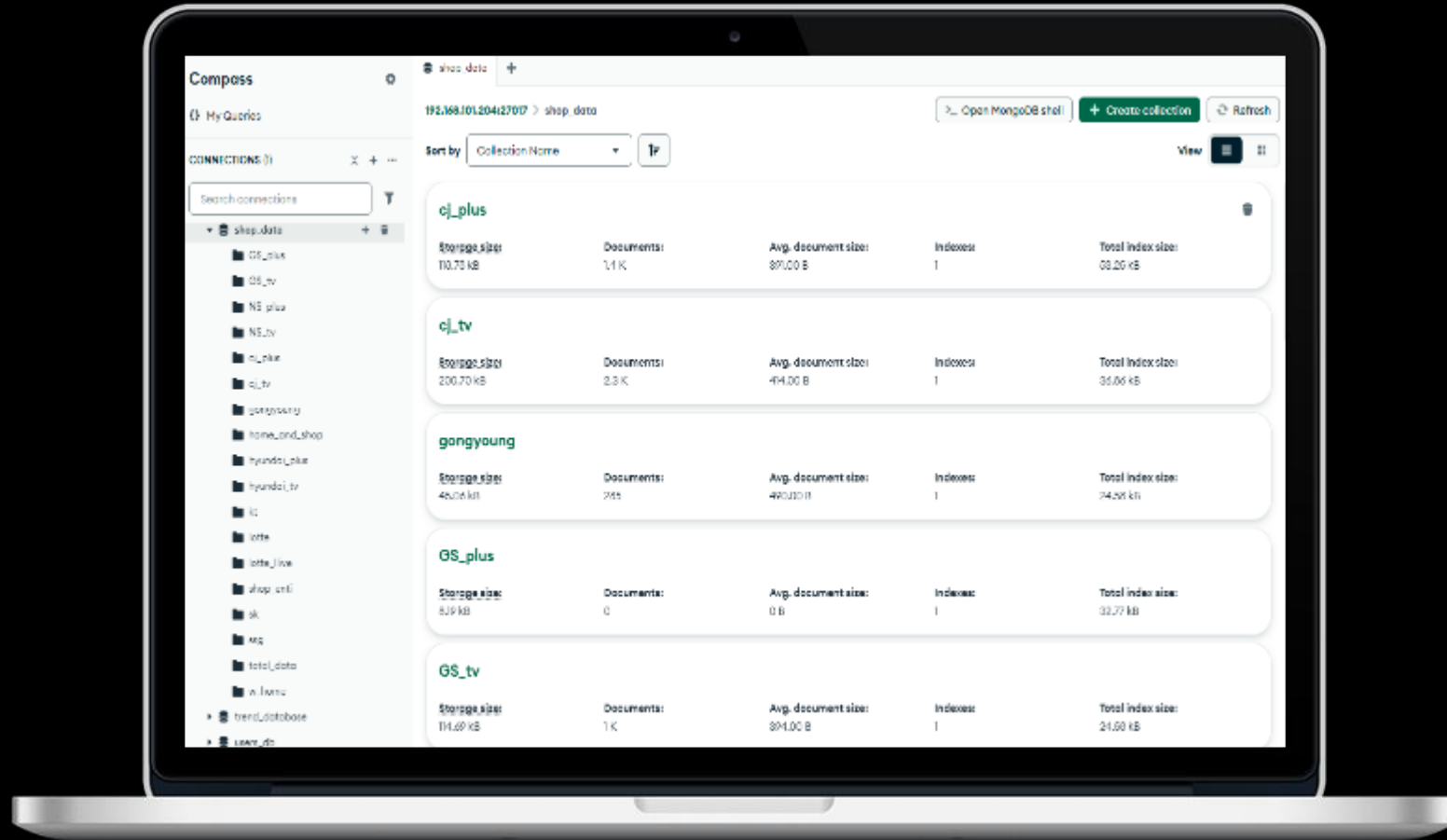
Keyword1	Keyword2	Keyword3
가슴기	가열식	공기청정
조끼	특가	할인
연말세일	특가	할인혜택
베이비용품	육아용품	크리스마스 선물
유아용품	안전매트	육아용품

키워드



3. 전처리

데이터 베이스 – Mongo DB

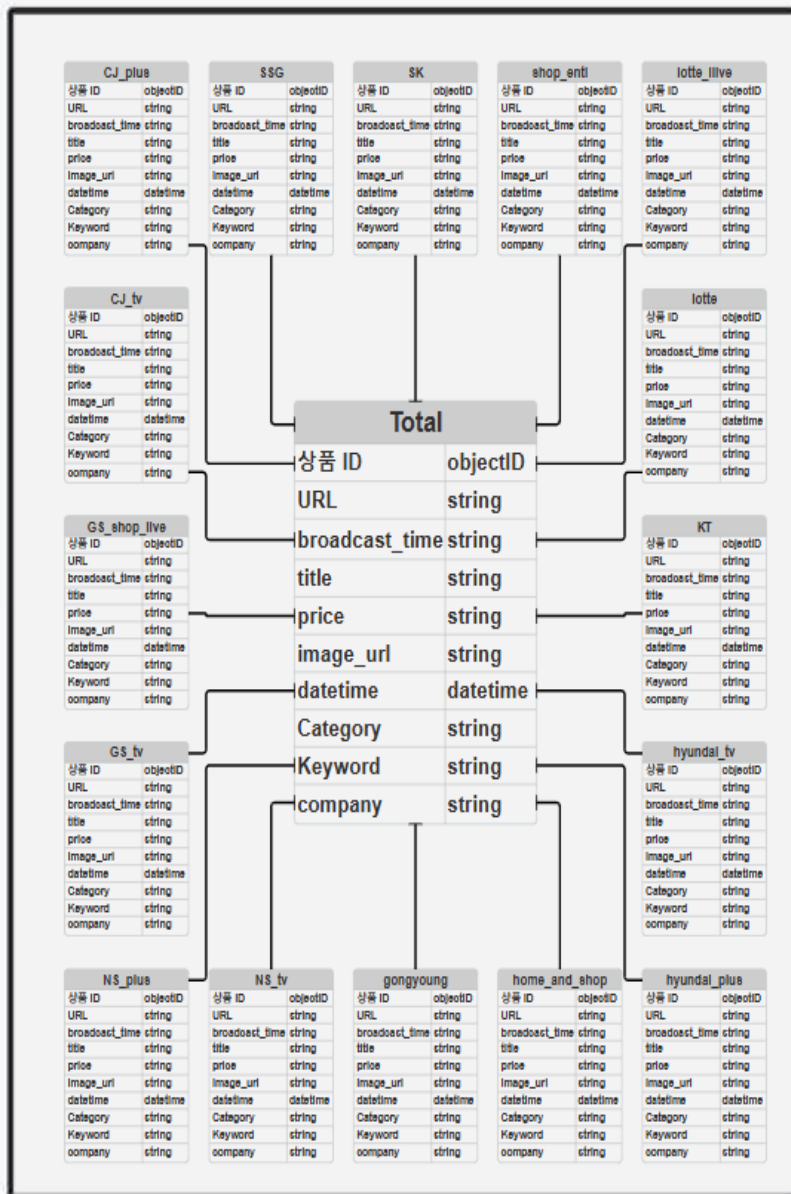


4. 데이터 베이스 설계

데이터 베이스 설계 ERD

Users_db	
_id	objectID
username	datetime
nickname	string
name	string
phone	string
email	string
password	string
birthday	string
gender	string
preferences	string

NAVER	
_id	objectID
datetime	datetime
categories	string
Keyword	string
rank	string
period	string



4. 모델링

1. 데이터
2. FastText
3. Clustering
4. Auto Encoder

1. 데이터

가상의 유저
'dydtns76'

dydtns76의 로그 데이터

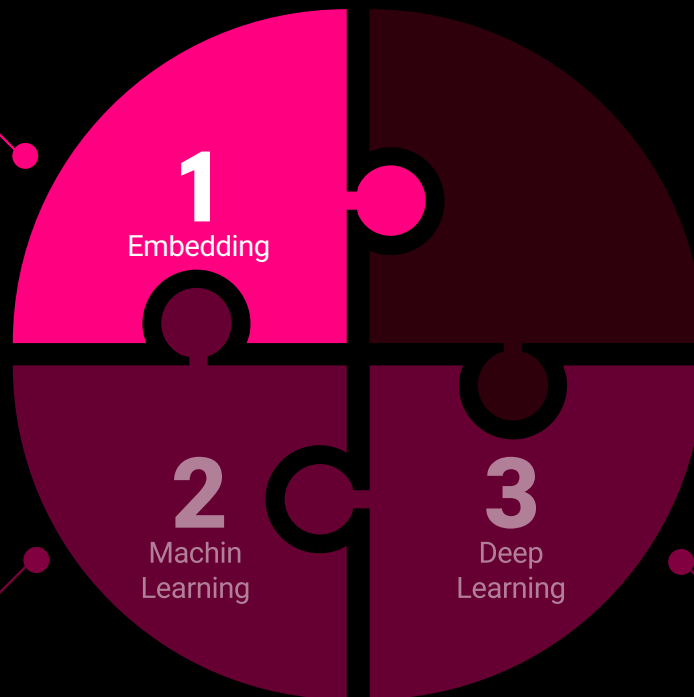
user_id	age	gender	preferred_category
dydtns76	20대	여성	[패션의류, 디지털/가전, 패션잡화]

user_id	이벤트	이벤트 시간	이벤트 내용
dydtns76	즐거찾기 추가	2025-01-01 23:12:11	677366e6d32e6f3e277cf68a
dydtns76	클릭	2025-01-01 23:20:00	677366e6d32e6f3e277cdcfa
dydtns76	클릭	2025-01-01 23:19:50	677366e6d32e6f3e277cdd0d
dydtns76	즐거찾기 추가	2025-01-01 23:15:00	677366e6d32e6f3e277cdd12
dydtns76	검색	2025-01-01 23:14:00	코트
dydtns76	클릭	2025-01-01 23:12:11	677366e6d32e6f3e277cf68a
dydtns76	클릭	2025-01-01 23:10:20	677366e6d32e6f3e277cff72
dydtns76	검색	2025-01-01 23:05:00	여성 코트
dydtns76	클릭	2025-01-01 20:20:00	677366e6d32e6f3e277cec9b
dydtns76	클릭	2025-01-01 20:00:00	677366e6d32e6f3e277cec9a
dydtns76	클릭	2025-01-01 15:10:00	677366e6d32e6f3e277cdd8d
dydtns76	클릭	2025-01-01 15:00:00	677366e6d32e6f3e277cdd7e
dydtns76	검색	2025-01-01 23:14:00	목걸이
dydtns76	즐거찾기 추가	2025-01-01 08:13:12	677366e6d32e6f3e277ce940
dydtns76	클릭	2025-01-01 08:12:11	677366e6d32e6f3e277cf7ac
dydtns76	클릭	2025-01-01 08:10:20	677366e6d32e6f3e277ce946
dydtns76	검색	2025-01-01 08:00:00	냉장고

1. FastText

FastText

한국어 Wikipedia 데이터를 기반으로 훈련된
FastText 임베딩 모델 사용



Clustering

Auto Encoder

1. FastText

Embedding

✓ Embedding

상품 추천 서비스에서 유사도를 활용하려면 임베딩을 통해 데이터의 잠재적 특징을 벡터화 하고, 코사인 유사도 등으로 효율적인 추천을 구현하는 것이 유리

✓ 사용 모델

한국어에 사용하기 좋은 여러 후보군을 선정.

S-Bert, Klue-Bert, FastText, Word2Vec 중 가장 임베딩이 잘 되었던 FastText 모델을 채택

✓ FastText

단어를 문자 n-그램의 벡터로 학습하여 어휘 밖 단어(OOV) 처리와 언어 모델링을 효율적으로 수행하는 빠르고 확장성 높은 워드 임베딩 기법

Train

위키피디아데이터 학습

위키피디아 api를 사용하여
한국어 문장을 학습



Test

유사도 테스트

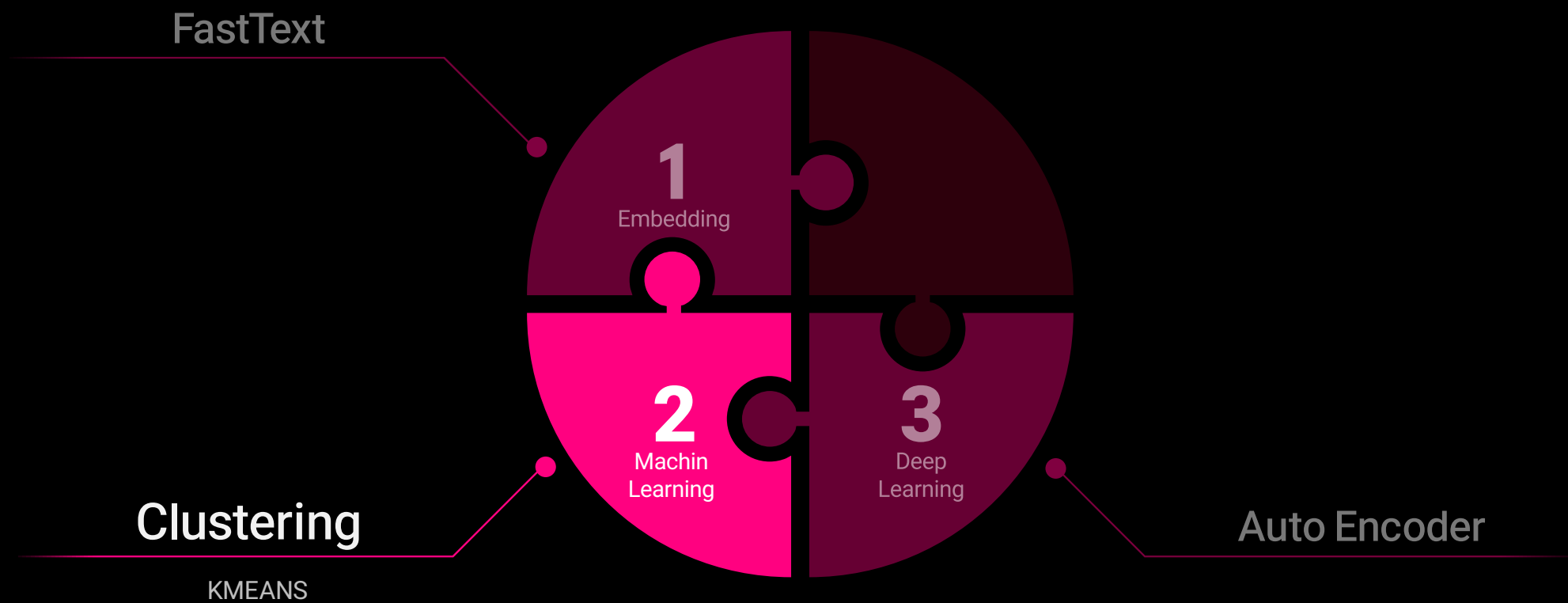
실제 홈쇼핑 title을 넣어, 상품간의 유사도를
확인하여 임베딩이 제대로 이루어졌는지 확인.



Model

fastText

3. Clustering



군집화 분석 과정

1. 입력 데이터

- ❖ Open Ai 로 추출한 카테고리 및 3개의 키워드 컬럼



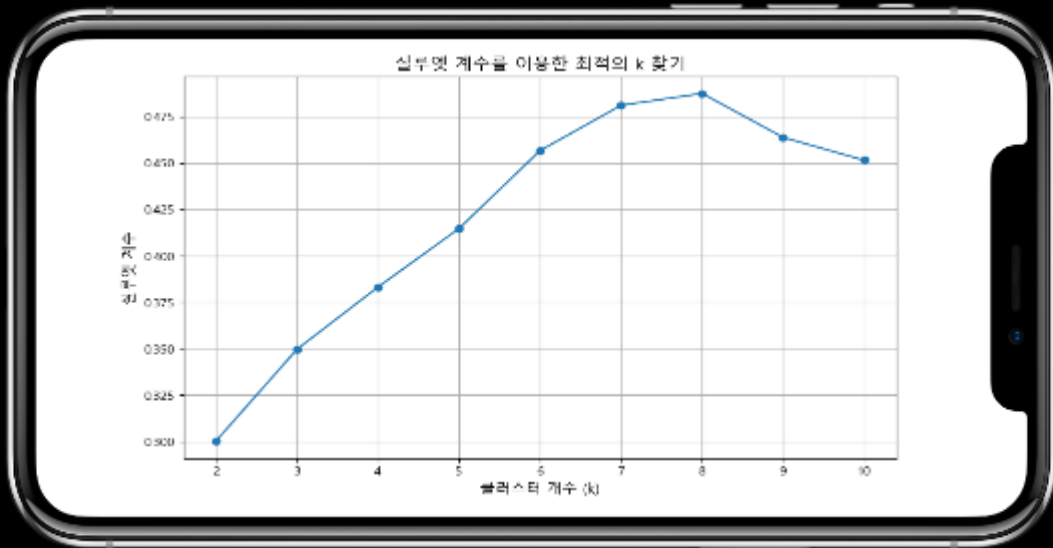
2. 임베딩

- ❖ 키워드 및 카테고리 컬럼을 하나의 컬럼으로 생성 후 FastText 모델로 임베딩



3. 군집화

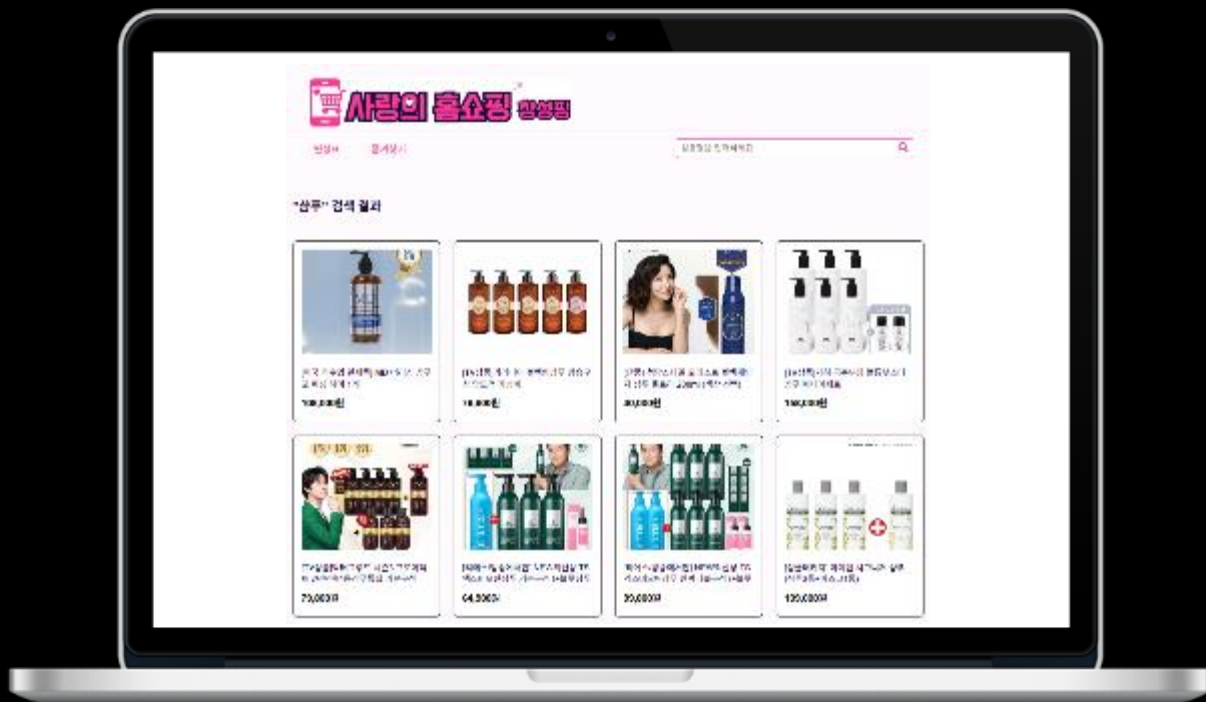
- ❖ 임베딩된 벡터를 평균값을 계산해 군집화



- ✓ 상품의 카테고리나 특성에 따라 유연하게 클러스터 수를 조정할 수 있어, 다양한 상품군에 맞는 추천 가능
- ✓ 사용자가 관심있을 만한 다른 상품을 추천함으로써 사용자에게 의미 있는 추천 가능

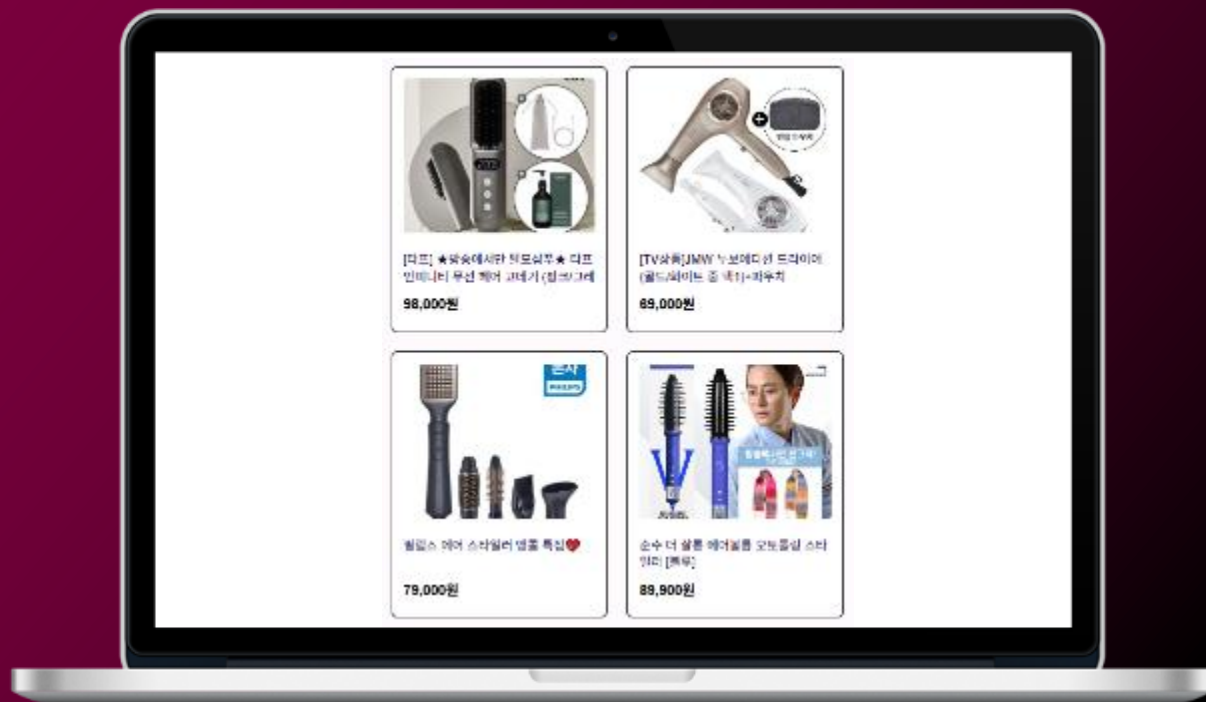
3. Clustering

검색 결과



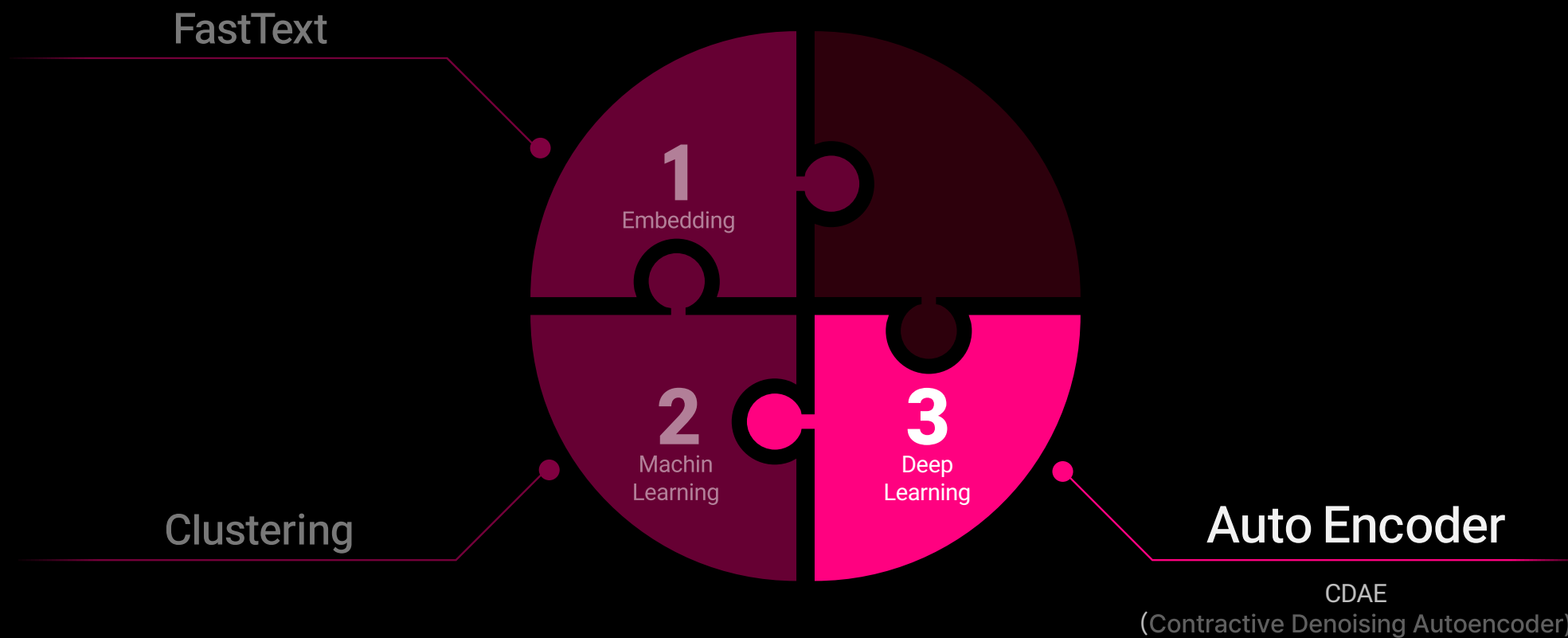
‘샴푸’ 조회 결과

연관 상품 추천



‘고데기, ‘드라이기’ 등 연관 상품 추천

4. Auto Encoder



Contractive Denoising Autoencoder

모델 설명

CDAE는 입력 데이터의 노이즈를 제거하면서 입력의 중요한 특징을 학습하고, 추가로 데이터의 작은 변화에 둔감하게 만들기 위해 입력의 변화에 대한 출력의 민감도를 줄이는 정규화를 적용한 자동 인코더.

“개인 맞춤 추천 서비스 모델

아키텍처

- Encoder
상품 데이터를 저차원의 벡터로 전환
Gaussian Noise와 Dropout으로
일반화
- Decoder
결합된 정보를 디코딩하여 각 상품의 추천
점수 계산

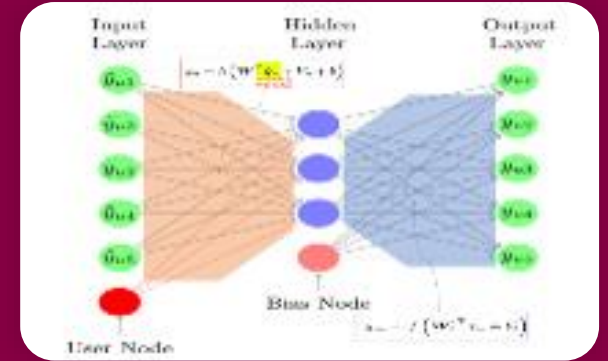
가중치

이벤트 행위에 따라 유사 제품 가중치

로그 데이터 발생 시간에 따른 가중치

트렌드 키워드 포함시 가중치

사용자 선호 카테고리 가중치



4. Auto Encoder

유저 데이터

user_id	age	gender	preferred_category
dydtns76	20대	여성	[패션의류, 디지털/가전, 패션잡화]

선호 카테고리
동일 상품 가중치

로그 데이터

user_id	이벤트	이벤트 시간	이벤트 내용
dydtns76	즐거찾기 추가	2025-01-01 23:12:11	677366e6d32e6f3e277cf68a
dydtns76	클릭	2025-01-01 23:20:00	677366e6d32e6f3e277cdcf
dydtns76	클릭	2025-01-01 23:19:50	677366e6d32e6f3e277cdd0d
dydtns76	즐거찾기 추가	2025-01-01 23:15:00	677366e6d32e6f3e277cdd12
dydtns76	검색	2025-01-01 23:14:00	코트
dydtns76	클릭	2025-01-01 23:12:11	677366e6d32e6f3e277cf68a
dydtns76	클릭	2025-01-01 23:10:20	677366e6d32e6f3e277cff72
dydtns76	검색	2025-01-01 23:05:00	여성 코트
dydtns76	클릭	2025-01-01 20:20:00	677366e6d32e6f3e277cec9b
dydtns76	클릭	2025-01-01 20:00:00	677366e6d32e6f3e277cec9a
dydtns76	클릭	2025-01-01 15:10:00	677366e6d32e6f3e277cdd8d
dydtns76	클릭	2025-01-01 15:00:00	677366e6d32e6f3e277cdd7e
dydtns76	검색	2025-01-01 23:14:00	목걸이
dydtns76	즐거찾기 추가	2025-01-01 08:13:12	677366e6d32e6f3e277ce940
dydtns76	클릭	2025-01-01 08:12:11	677366e6d32e6f3e277cf7ac
dydtns76	클릭	2025-01-01 08:10:20	677366e6d32e6f3e277ce946
dydtns76	검색	2025-01-01 08:00:00	냉장고

최근에 이벤트가 발생한
상품일수록 유사 상품에 가중치

즐거찾기 > 검색 > 클릭

해당 날짜 기준 패션의류 트렌드 키워드 1위는 '패딩'

사용자가 자주 클릭한 '코트'와 트렌드 키워드인 '패딩'이 함께 적용된 모습

모델 총평



❖트렌드를 적용해서 잘 추천해주는 것을 볼 수 있음.
개인의 맞춤에 따라 가중치를 다르게 설정하여 개인의 맞춤에 맞는 모델을 만들 수 있을 것으로 예상됨.
ex) 트렌드 기반으로 추천 받기 싫은 고객을 위해 트렌드 가중치를 제거.

추천상품 Top5

상품 ID: 677366e6d32e6f3e277cf534
제목: 험머 남녀 롱패딩코트
카테고리: 패션의류
가격: 59000.00
추천 점수: 0.9982

상품 ID: 677366e6d32e6f3e277cec8b
제목: [TV상품]머텔 공용 24FW 최신상
카테고리: 패션의류
가격: 0.00
추천 점수: 0.9980

상품 ID: 677366e6d32e6f3e277d22f0
제목: [기획초특가] 릴리전 에어 롱패딩
카테고리: 패션의류
가격: 59000.00
추천 점수: 0.9978

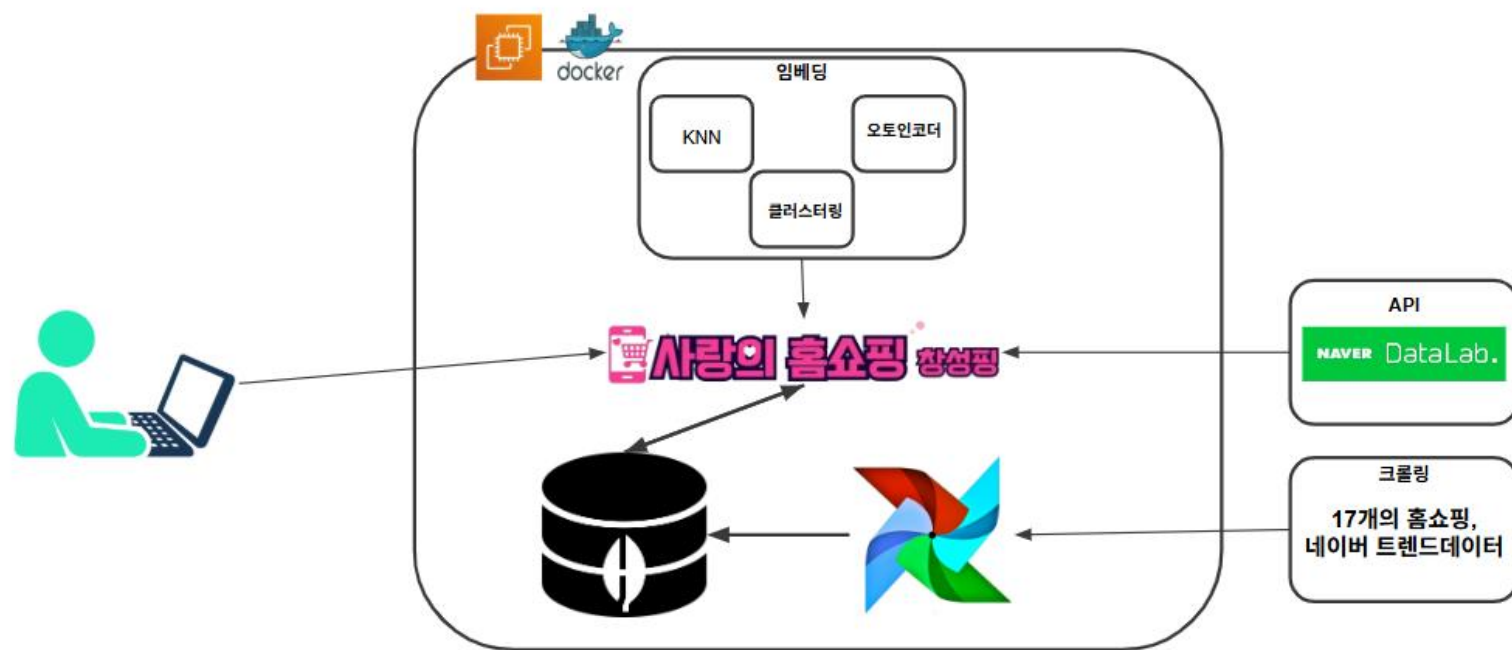
상품 ID: 677366e6d32e6f3e277d1ea9
제목: 24FW 최신상 로베디카파 롱패딩
카테고리: 패션의류
가격: 99000.00
추천 점수: 0.9975

상품 ID: 677366e6d32e6f3e277d237d
제목: [기획초특가] 릴리전 에어 롱패딩
카테고리: 패션의류
가격: 59000.00
추천 점수: 0.9973

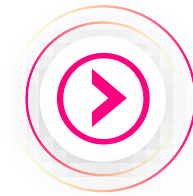
5. WEB

1. 설계.
2. 시연영상

1. 설계



2. 시연 영상



6. 결론

1. 향후개선방향
2. Q&A

1. 향후 개선 방향

1. 동일한 상품 중복



- 방송 title이 다르나 같은 상품을 판매하는 홈쇼핑이 있음. 이에 대해 대비가 되어있지 않음.
- 상품 구성을 달리해 판매하는 상품에 대한 개선 필요.

2. 사용자 데이터



- 사용자 데이터가 없어 가상의 사용자를 만들어 가상의 로그데이터를 사용.
- 추후 실제 데이터를 활용해 알고리즘을 개선.

3. 데이터 부족



- 일정 상품에 대한 데이터가 부족해, 위키피디아 API를 활용해 임베딩을 진행했으나, 향후 크롤링 자동화로 데이터가 점차 많아질 것으로 예상

4. 상품 페이지 추가



- 상품을 클릭했을 때, 해당 홈쇼핑 사이트로 이동하는 것이 아닌, 상품 페이지를 추가하여 해당 페이지로 이동할 수 있도록 수정.
- 상품 페이지에서 연관 상품을 추천할 수 있도록 개선.

Question & Answer

사랑의 홈쇼핑 창성핑

#홈쇼핑
#상품 추천
#트렌드
#서비스



An aerial photograph of a city skyline, featuring numerous skyscrapers and dense urban development. The image is overlaid with a semi-transparent purple filter. The text "“감사합니다.”" is centered in the upper half of the image in a white, bold, sans-serif font.

“감사합니다.”