



# AiTEMS 쇼핑추천의 시작부터 HyperScale AI 도입까지

양민철 NAVER  
김동현 NAVER

# CONTENTS

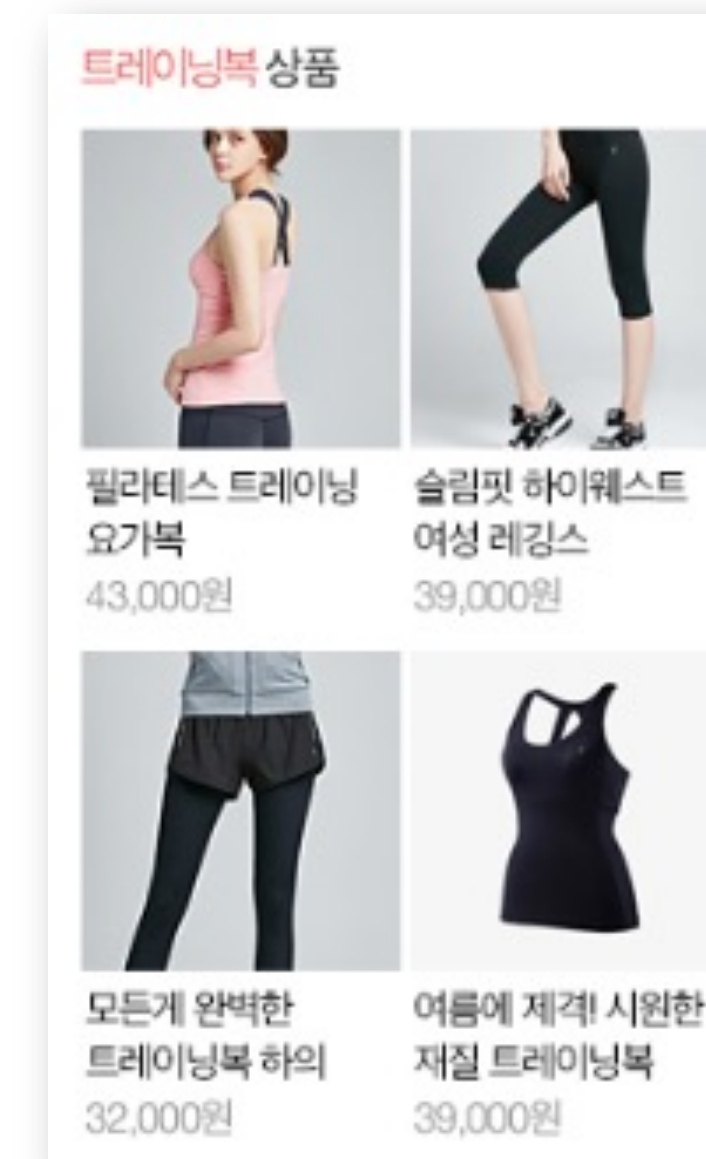
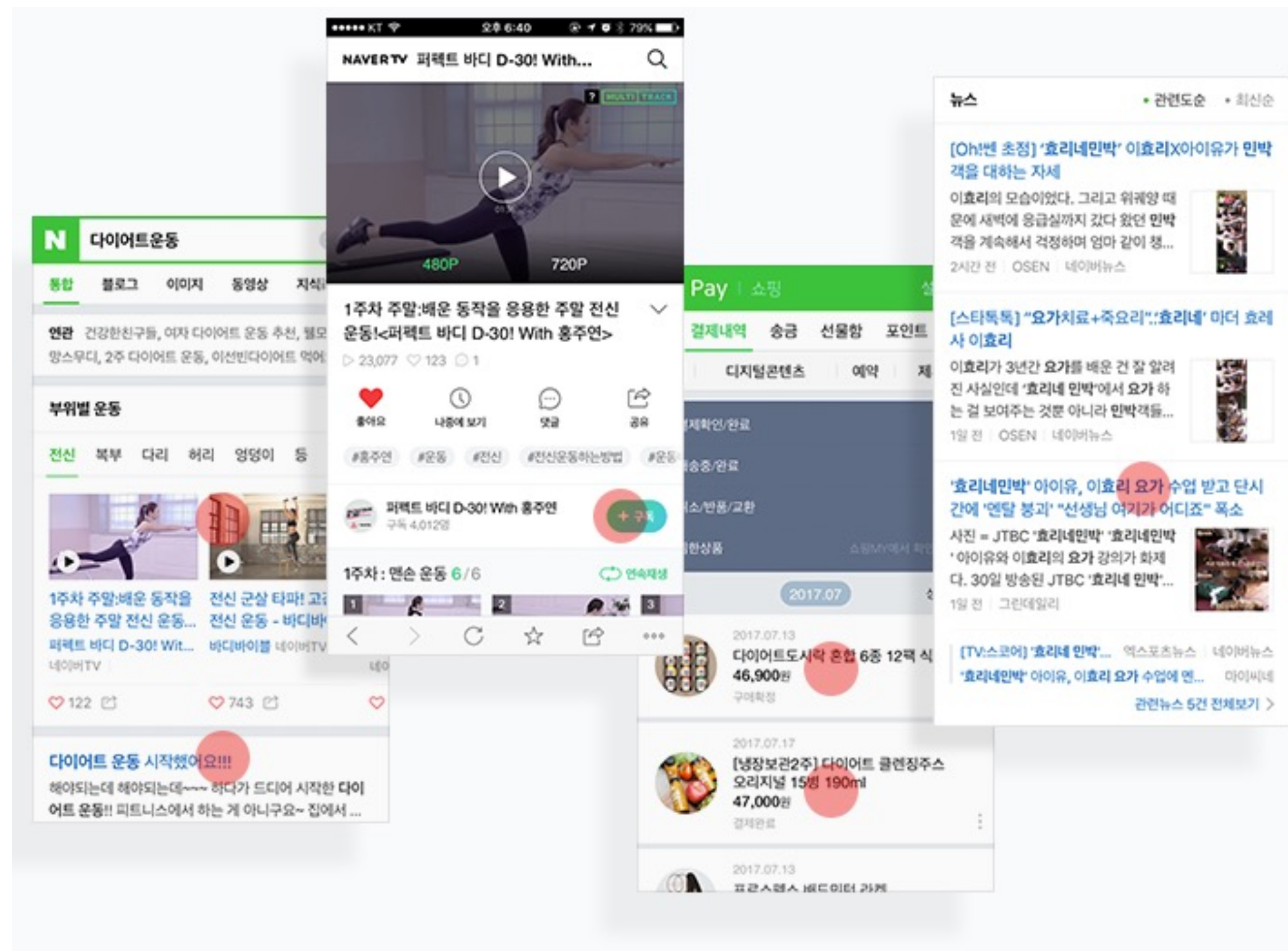
1. 네이버 쇼핑추천 AiTEMS 개요
2. 다양한 쇼핑 니즈를 만족시키기 위한 기술/서비스 개선 사례
3. AiTEMS X HyperCLOVA

# 1. 네이버 쇼핑추천 AiITEMS 개요

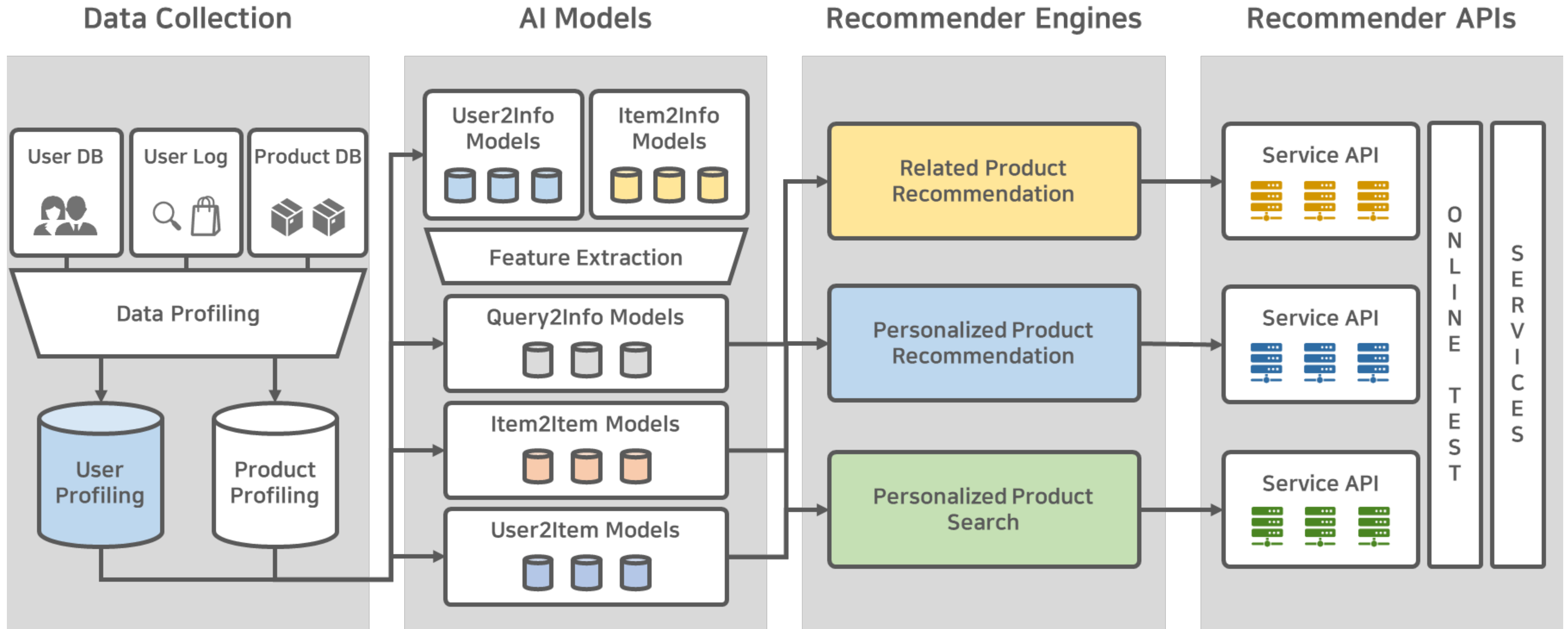
# 1.1 AiTEMS란?

**Ai + Items = AiTEMS** (에이아이템즈)

인공지능이 사용자 맞춤 **상품**을 추천



# 1.2 AiTEMS 플랫폼 형태로 진화



# 1.3 다양한 쇼핑추천 도입 중 @네이버쇼핑

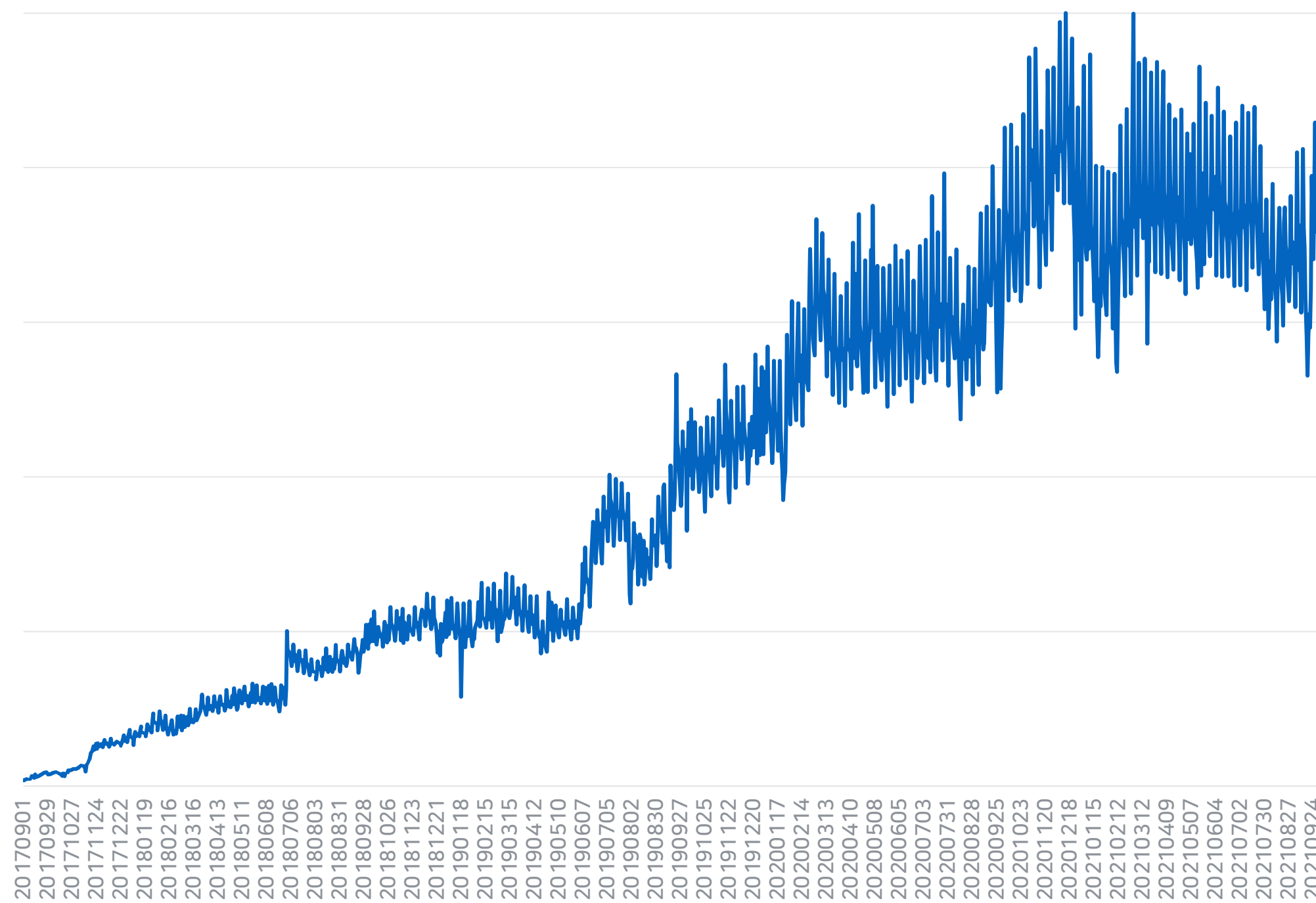
## AiTEMS 쇼핑추천 라인업

네이버메인	쇼핑검색	쇼핑서비스	주문완료
추천상품	내 관심 상품	연관 상품 추천	구매 후 추천
추천트렌드	반응형 추천	패션 스타일 추천	
재구매 추천	트렌드 검색	뷰티 리뷰 추천	
반응형 추천	선호몰 검색	펫 상품 추천	
내 관심 상품	개인화 랭킹		
	선물 추천		

# 1.4 쇼핑추천을 통한 소비는 상승 중

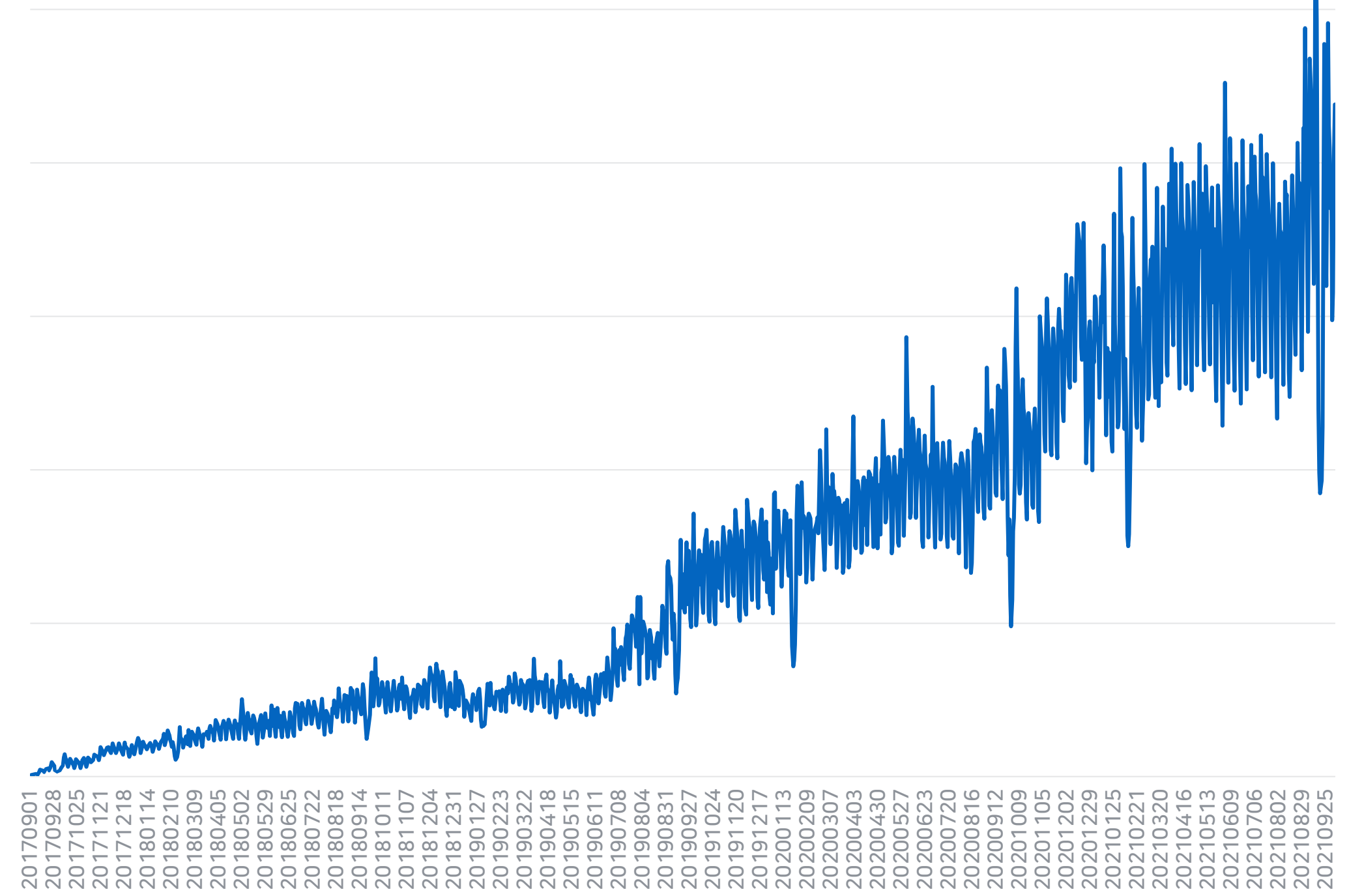
AiTEMS를 통한 상품 클릭 수

네이버쇼핑  
10% 점유



AiTEMS를 통한 상품 거래액

네이버쇼핑  
8% 점유

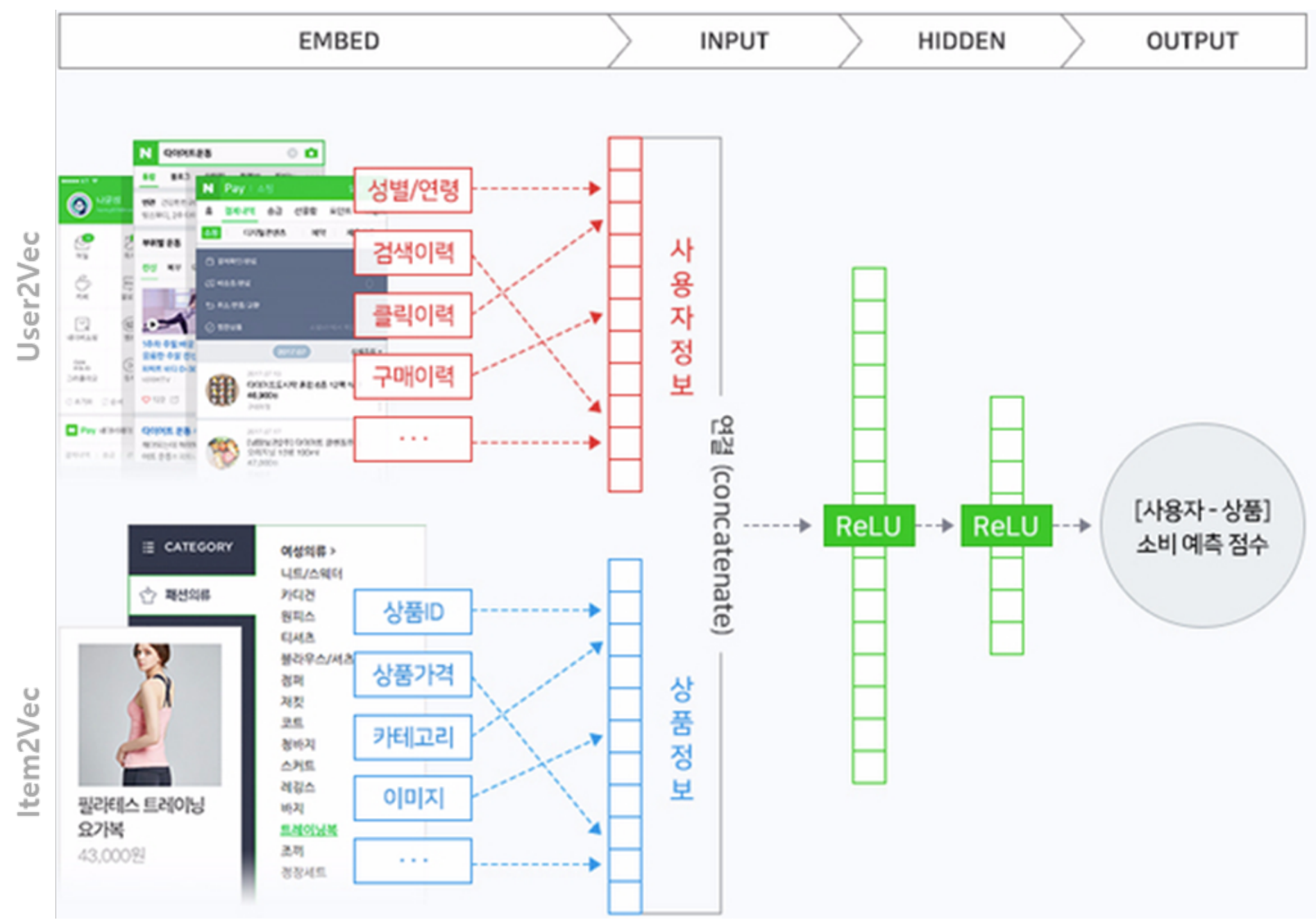


## 2. 다양한 쇼핑 니즈를 만족시키기 위한 기술/서비스 개선 사례



# 2.1 어떻게해야 사용자를 만족시킬수 있을까?

## [네이버 메인] 개인화 상품추천: 출시 초기



2017년 8월 31일  
**AiTEMS 추천상품 출시**  
 미지근한 사용자 반응

↓

딥러닝 기술을 믿었으나,  
 사용자의 공감을 왜 얻지 못했을까?

[참고 논문] P. Covington et al. (Google), *Deep Neural Networks for YouTube Recommendations*, RecSys 2016  
 [특허 출원] 딥러닝을 활용한 개인화 상품 추천 (출원 번호: 10-2017-0044834)

## 2.1 어떻게해야 사용자를 만족시킬수 있을까?

### [네이버 메인] 개인화 상품추천: 사용자 피드백 수집

“왜 저한테 이 상품들이 추천되나요?”

사용자가 공감할만한 **테마**가 필요하다.

“추천된 상품이 저와 **어울리지 않아요**”

**명확하고 구체적인 취향 분석**이 필요하다.



# 2.1 어떻게해야 사용자를 만족시킬수 있을까?

## [네이버 메인] 개인화 상품추천: 명확한 추천사유 수집

카테고리 기반 추천

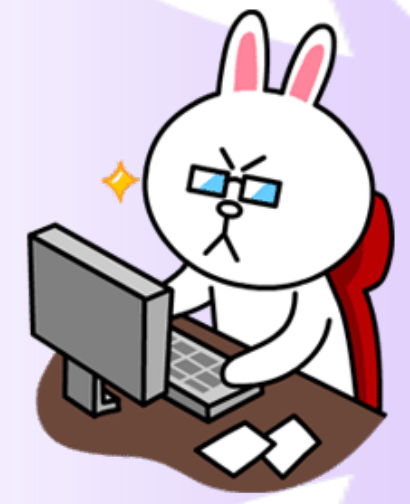
사용자가 선호하는  
브랜드, 쇼핑몰, 검색어 기반 추천



원피스



운동화



“왜” 추천되는지?  
세분화 된 추천

기모롱 원피스



임블리 원피스



겨울 운동화



나이키 운동화



# 2.1 어떻게해야 사용자를 만족시킬수 있을까?

## [네이버 메인] 개인화 상품추천: 명확한 추천사유 제공

사용자가 선호하는  
브랜드, 쇼핑몰, 검색어 기반 추천

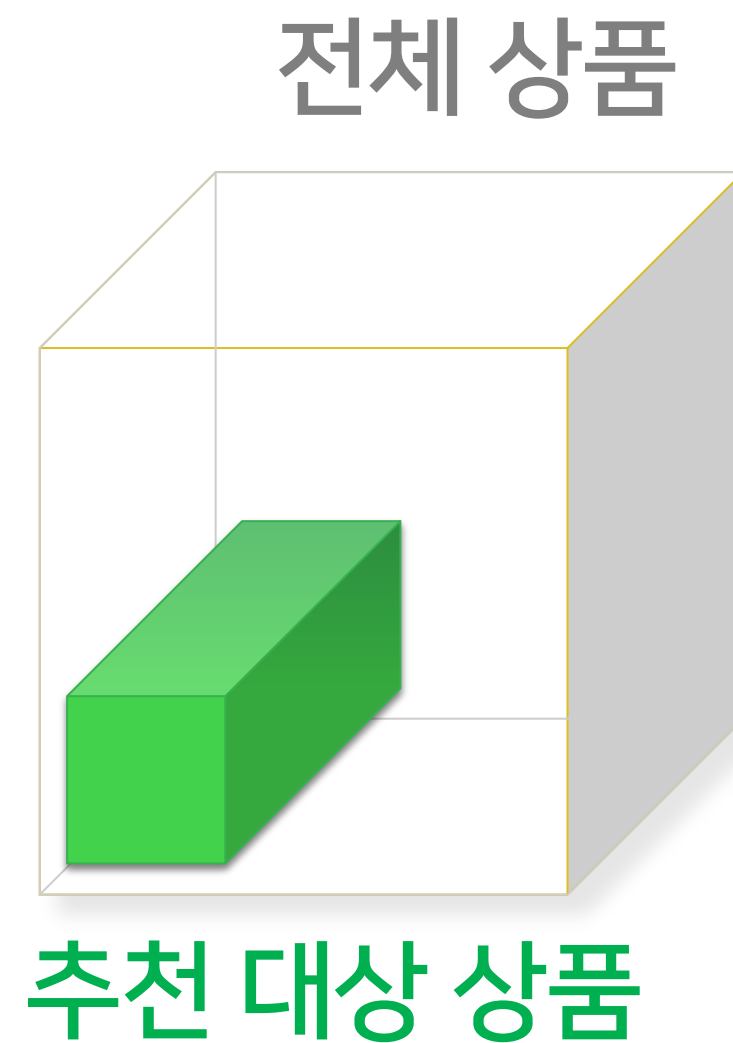


다양하고 세분화된  
추천 사유 제공 후

상품 클릭 8배 상승

## 2.2 Cold-Item 추천을 위한 모델 도입 사례

### [네이버 쇼핑] 인기 상품 Bias 문제



Popular-Item 치중된 추천

대중적인 선호도가 반영된 검색 결과 위주의 상품 소비

Cold-Item 활용한 추천 필요

자신만의 확고한 취향이 있는 사용자를 위한 추천  
(Personalization)

Cold-Item Recommendation

## 2.2 Cold-Item 추천을 위한 모델 도입 사례

### [기술] Item2Item: 상품 메타정보 (Side-Information) 활용



구매/클릭 여부와 관계 없는 정보 필요

모든 상품은 메타(컨텐츠) 정보를 가지고 있고,  
하나의 메타 정보는 다수의 상품이 공유

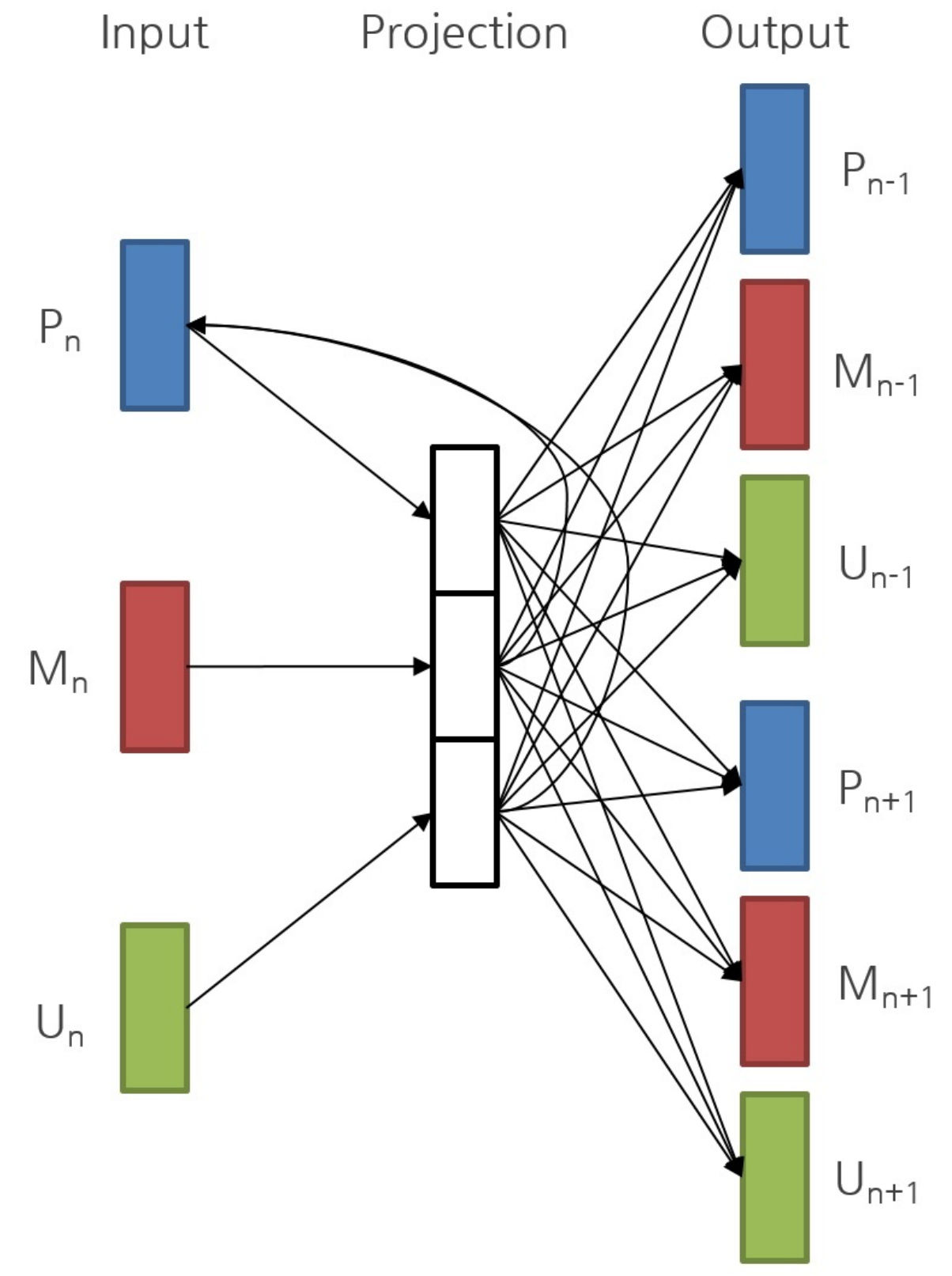
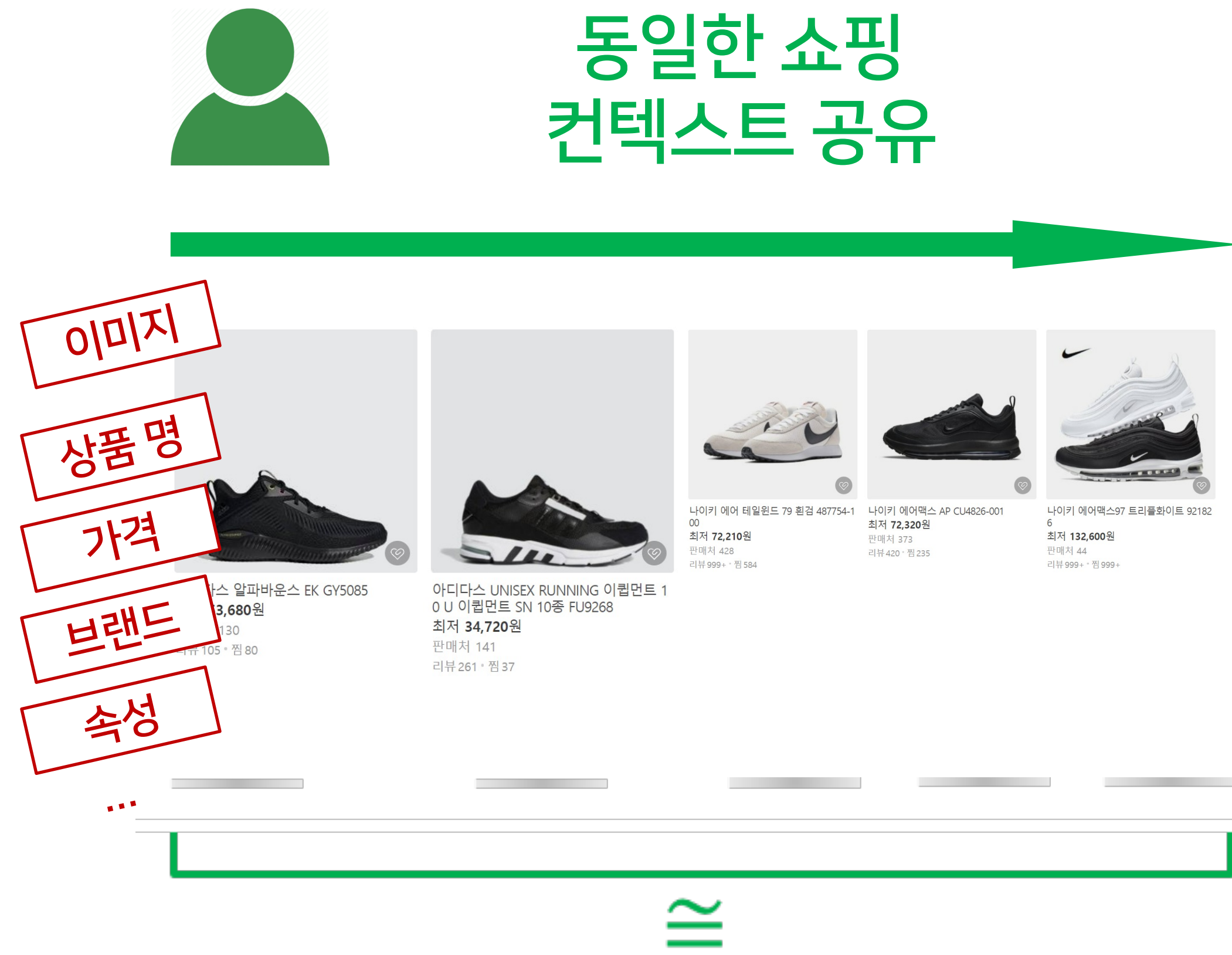
메타 정보를 통해 클릭 되지 않은 상품과  
클릭 된 상품의 관계를 찾아낼 수 있음

메타 정보 기반 상품 임베딩

# 2.2 Cold-Item 추천을 위한 모델 도입 사례

## [기술] Item2Item: 상품 메타정보 기반 학습 (Meta-Product Vector)

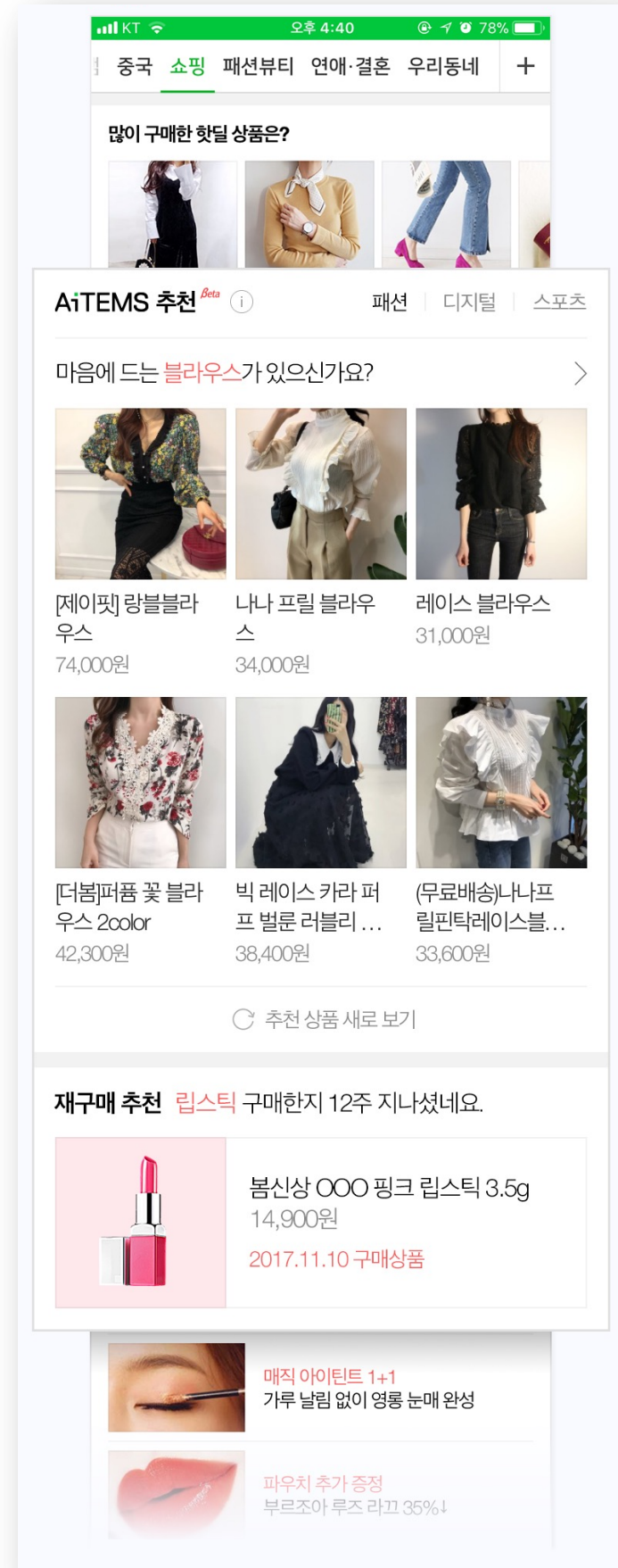
Meta-Product Vector



[참고 논문] Flavian Vasile et al., *Meta-Prod2Vec: Product Embeddings Using Side-Information for Recommendation*, RecSys 2016.  
 [특허 출원] 상품 추천 방법 및 이를 이용하는 서버 (출원 번호: 10-2018-0036900)

# 2.2 Cold-Item 추천을 위한 모델 도입 사례

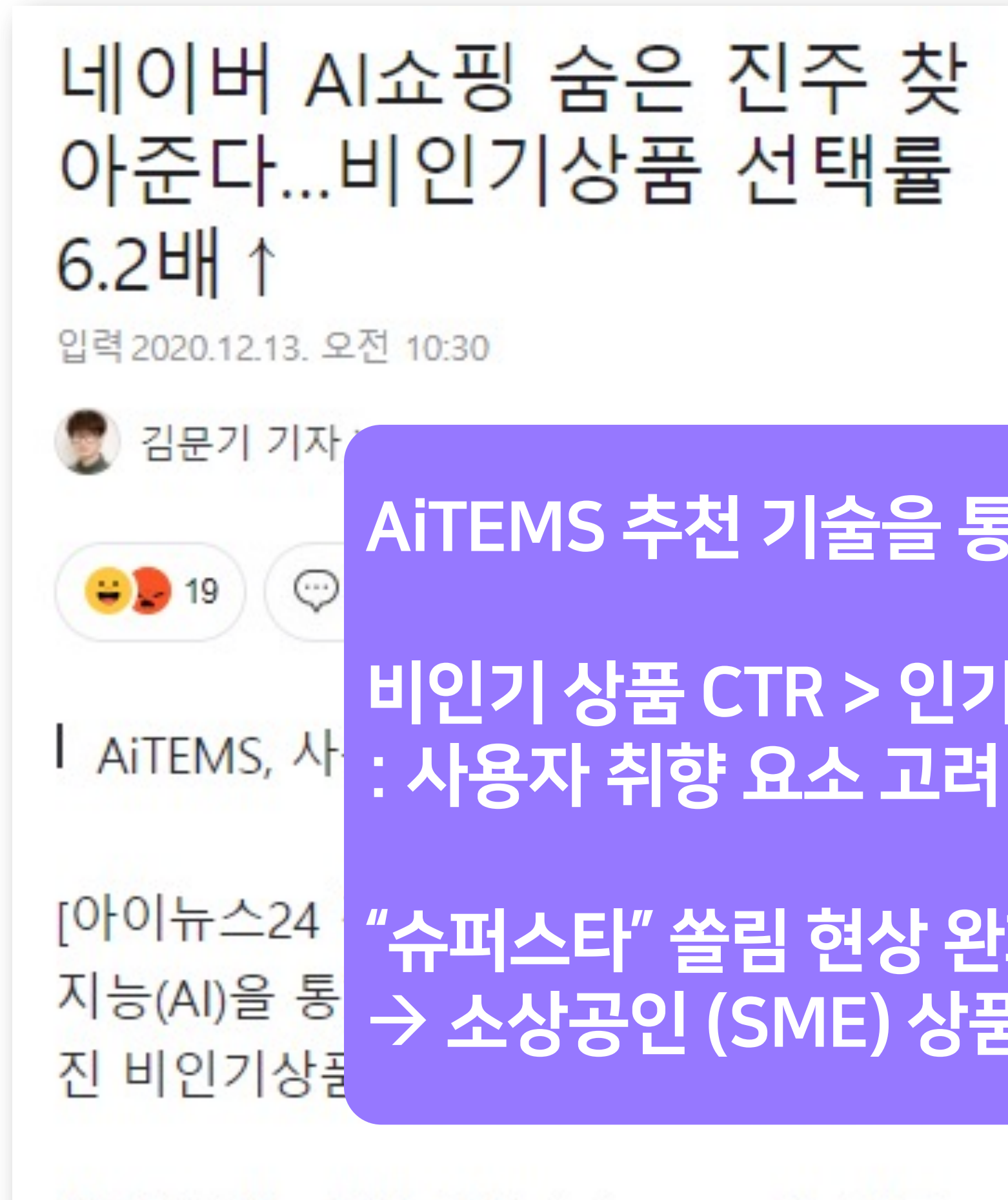
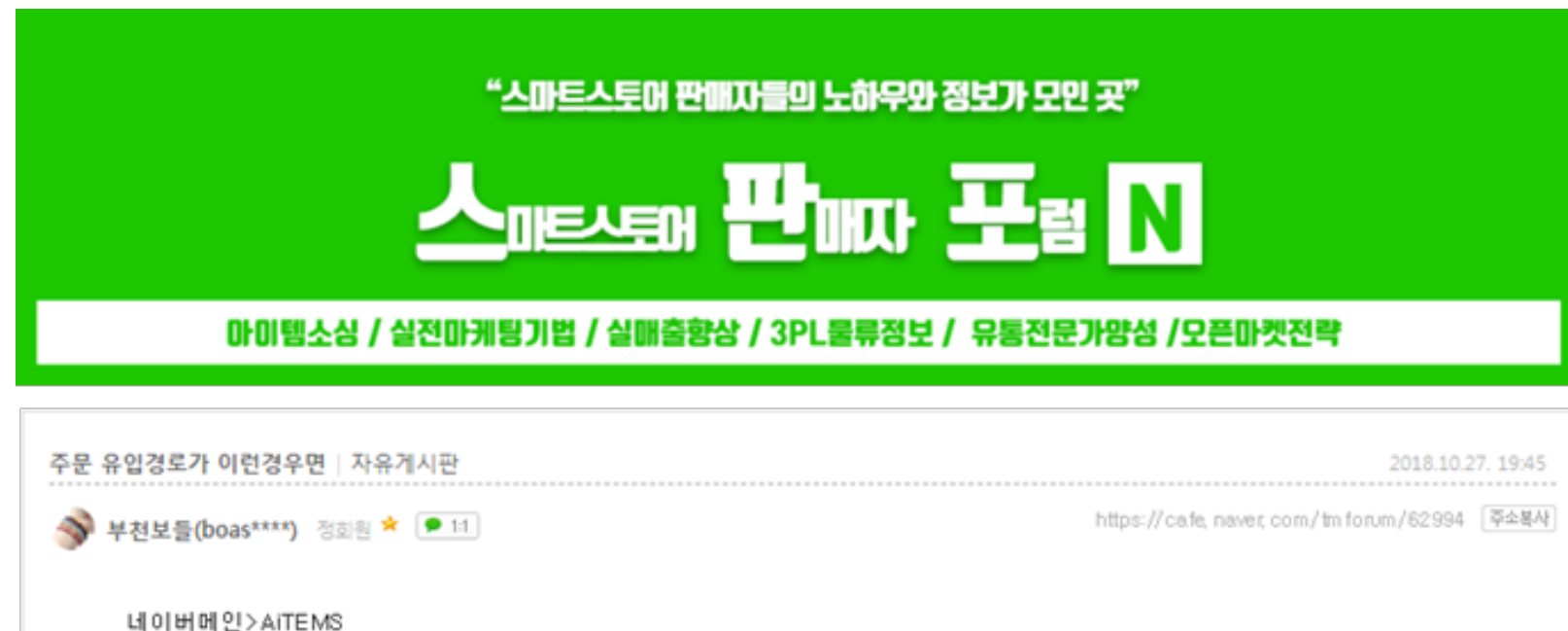
## [네이버 메인] 개인화 상품추천: Cold-Item / New-Item 추천 확장





# 2.2 Cold-Item 추천을 위한 모델 도입 사례

## [네이버 메인] 개인화 상품추천: 소상공인 도움이 되는 추천

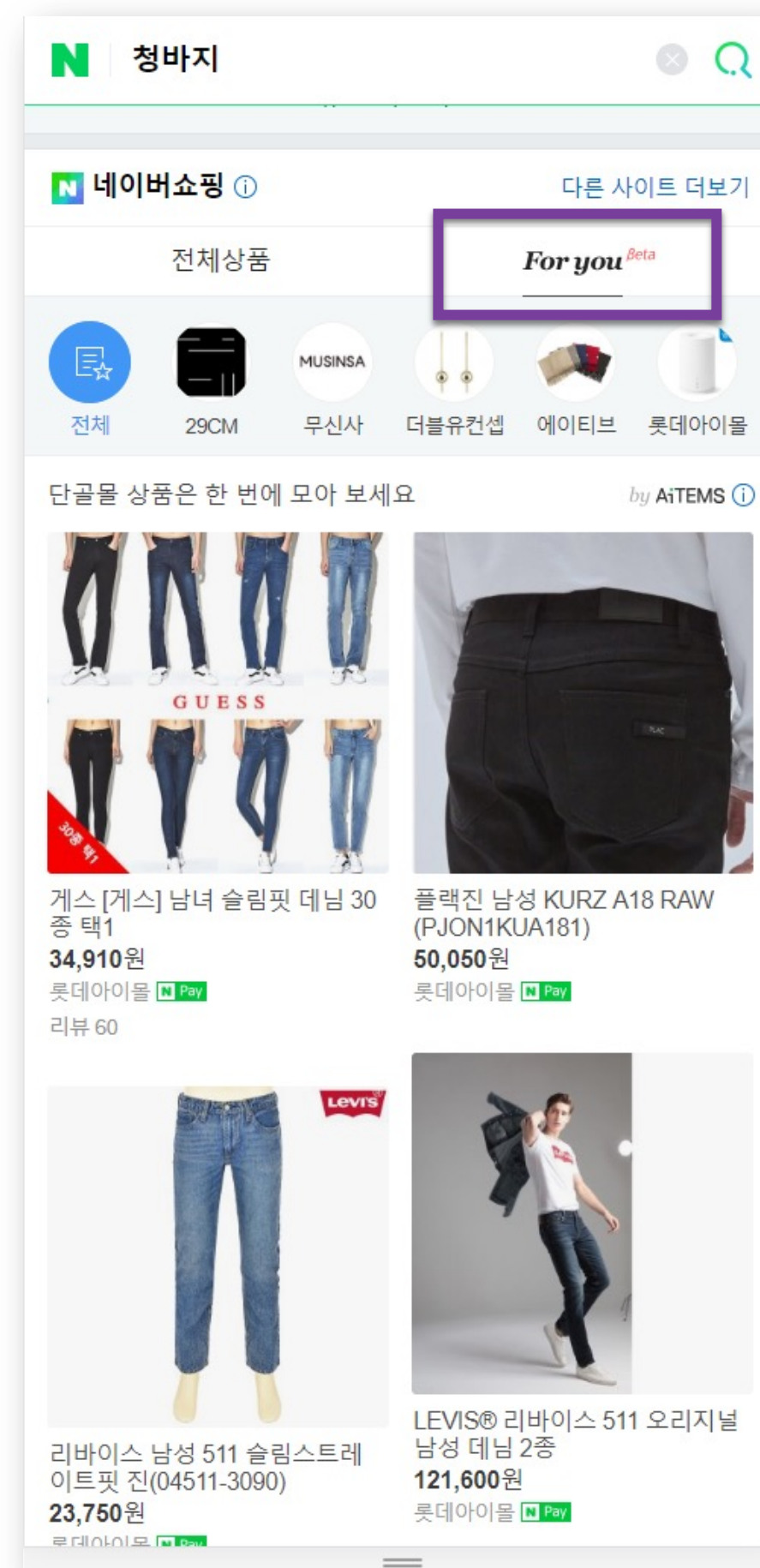


AiTEMS 추천 기술을 통해,  
 비인기 상품 CTR > 인기 상품 CTR  
 : 사용자 취향 요소 고려

“슈퍼스타” 쏠림 현상 완화  
 → 소상공인 (SME) 상품 판매 교두보

# 2.3 상품 검색 X 개인화 추천 도전기

## [쇼핑검색] For You 검색: 출시 초기



독립적인 탭 구조를 통해  
개인화 검색결과 제공



사용자들의 진입 장벽이 높음

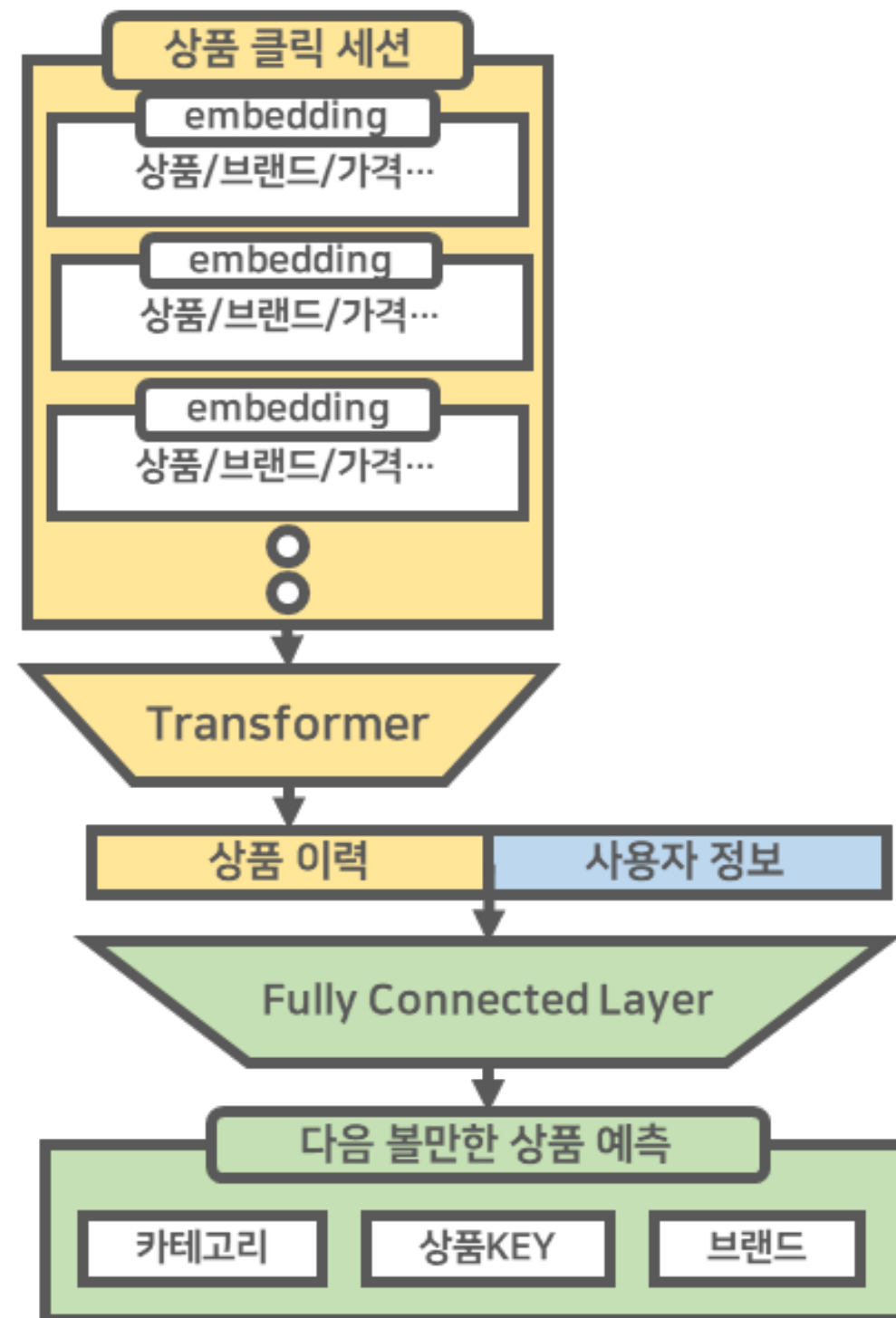


추가적인 사용자 액션 없이  
자연스러운 개인화 추천 제공 필요

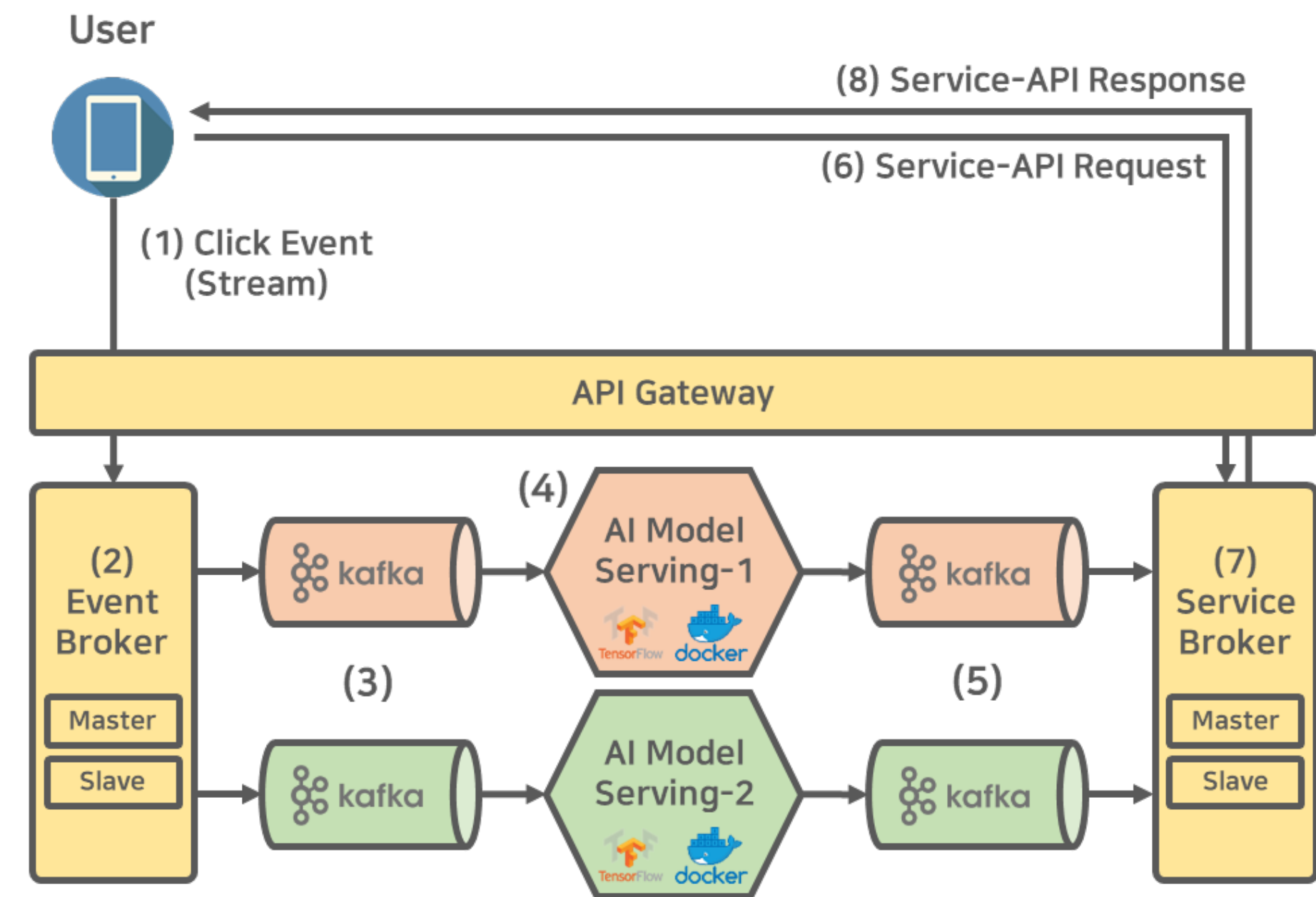
# 2.3 상품 검색 X 개인화 추천 도전기

## [기술] 세션 기반 추천: 실시간 추천 모델 & 서빙 엔진 도입

[모델] 취향 변화를 즉각 반영하는 추천



[시스템] 실시간 이력 수집 & 추천 결과 Inference

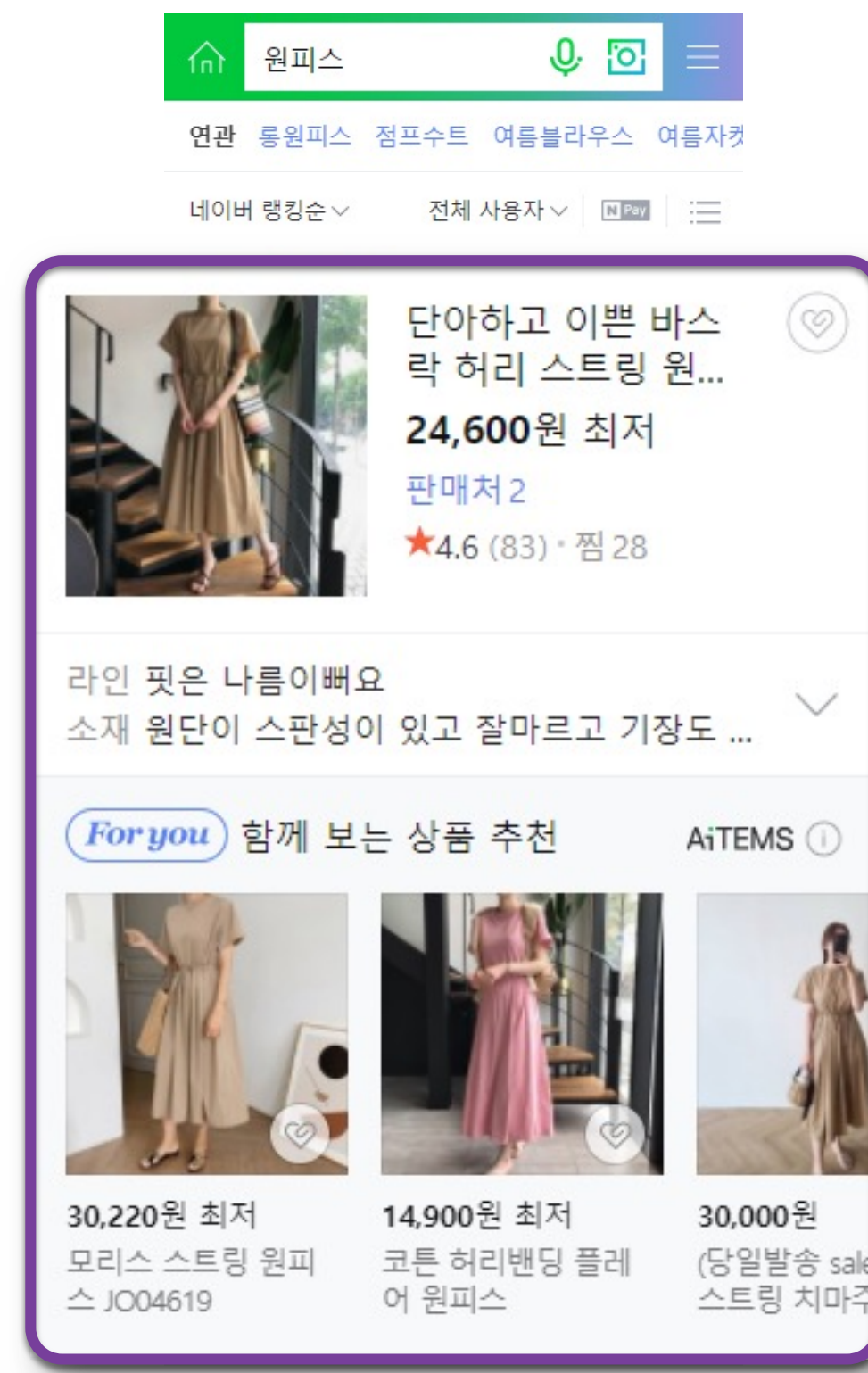
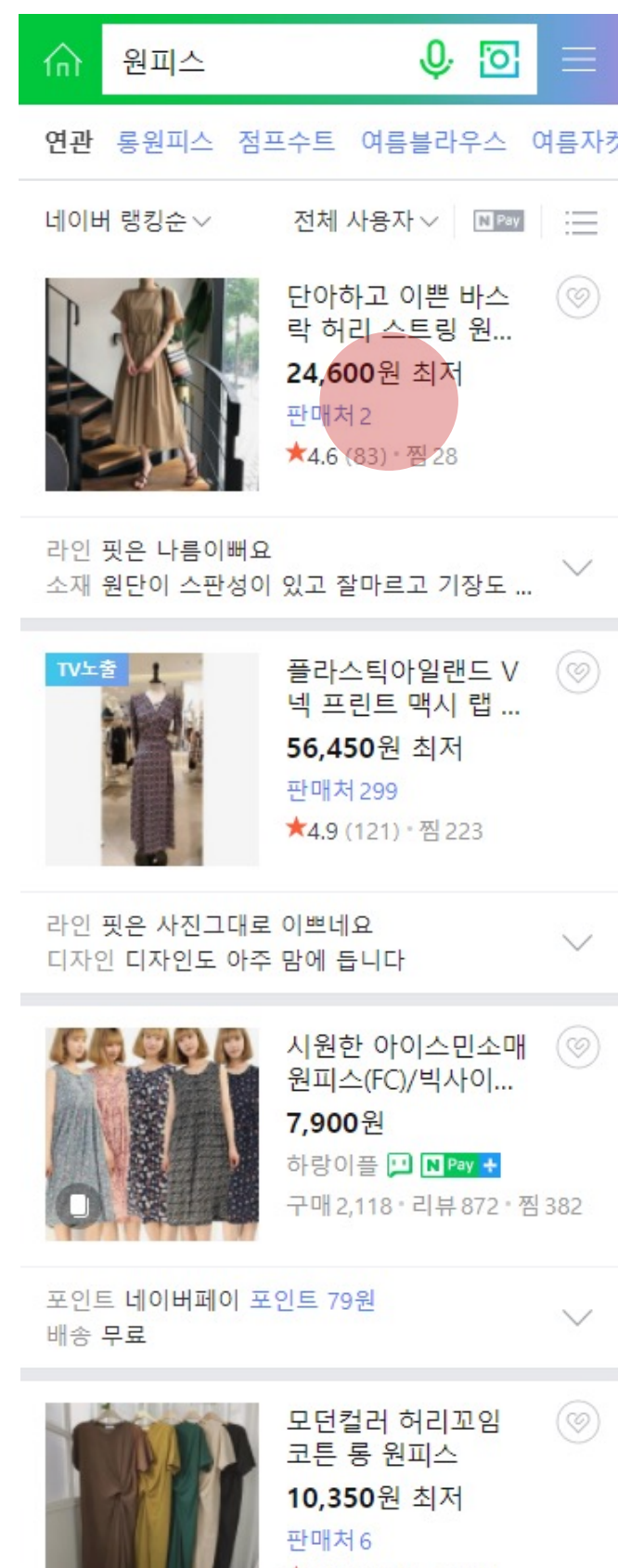


[참고 논문] A. Vaswani et al., Attention Is All You Need, NIPS 2017  
 [특허 출원] 반응형 쇼핑 상품 검색 (출원 번호: 10-2020-0027474)

# 2.3 상품 검색 X 개인화 추천 도전기

## [쇼핑검색] For You 검색: 반응형 추천 전환

검색 결과 상품 클릭 → 검색 결과로 돌아가기 → 상품 추천 영역 팝업



타 커머스 사이트 대비  
최초 반응형 추천 컨셉 도입

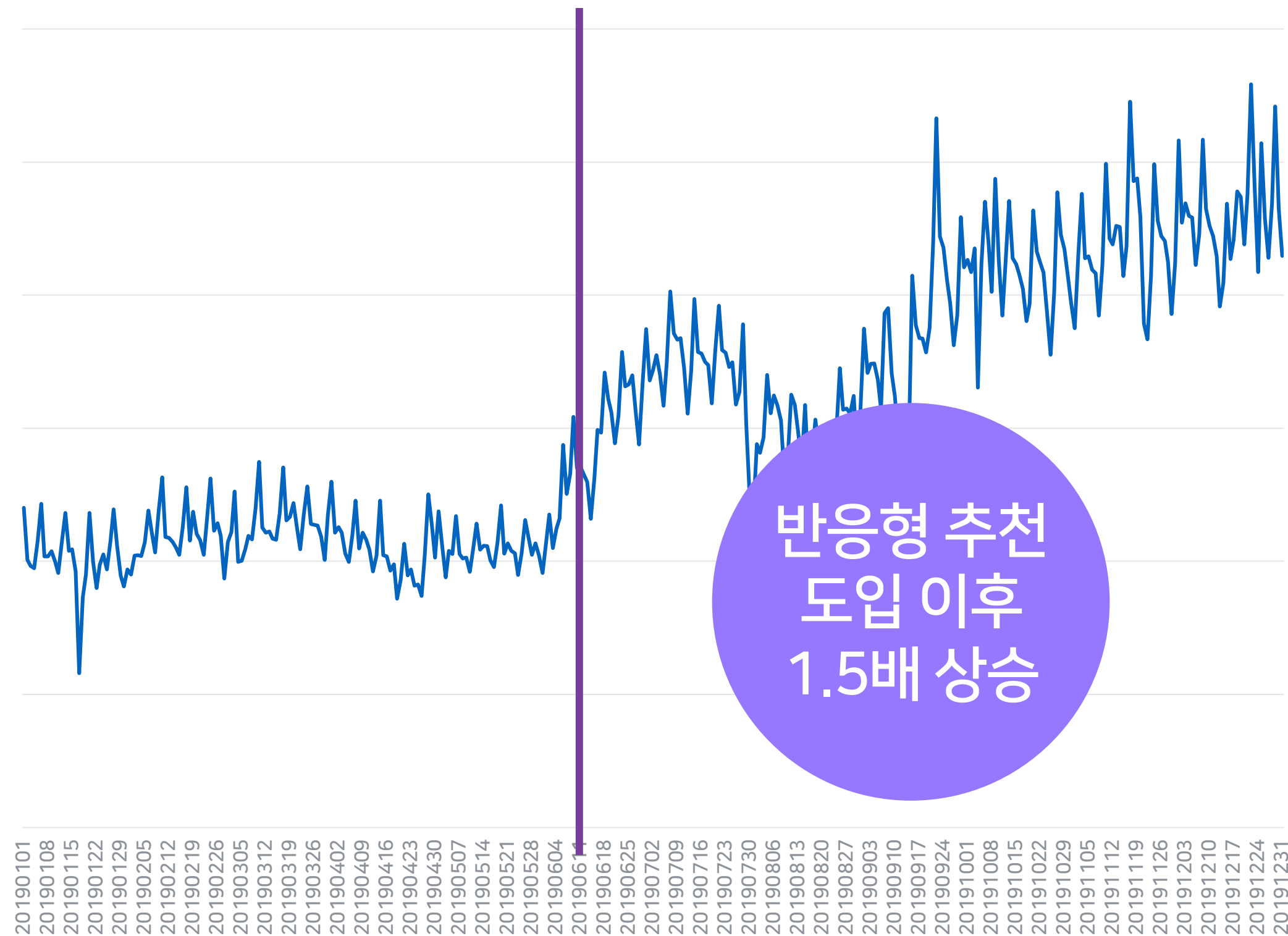
상품 탐색 과정에서 자연스럽게  
도움이 되는 추천 기술/서비스 설계

[특허 출원] 반응형 쇼핑 상품 검색 (출원 번호: 10-2020-0027474)

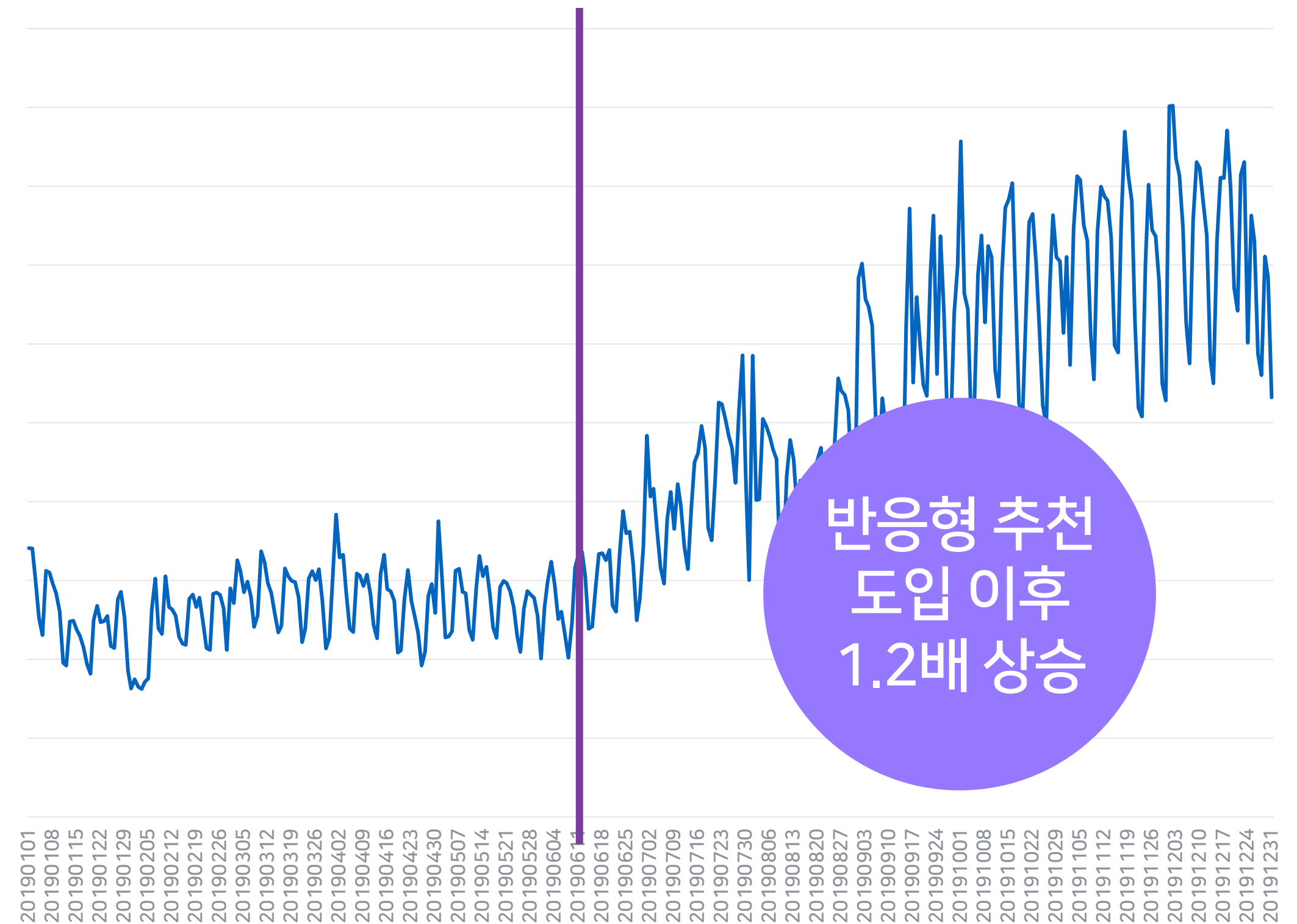
# 2.3 상품 검색 X 개인화 추천 도전기

## [쇼핑검색] For You 검색: 사용자 만족도 상승

AiTEMS를 통한 상품 클릭 수 (2019)



AiTEMS를 통한 상품 거래액 (2019)



## 2.4 추천 기술 X 서비스

개인화 쇼핑추천 기술 = Academy



Explainable Reco. (2017)

Cold-Start (2018)

Real-Time Reco. (2019)

... ?

“더 나은” 개인화 쇼핑추천 서비스

# 3. AiTEMS X HyperCLOVA

## 3.1 HyperCLOVA (KoreanGPT3)

**GPT3에 비해 한국어가 6500배 이상 학습된 한국어 초거대 언어모델 !**

- HyperCLOVA가 단순히 데이터만 다른 것은 아니니,
- 모델의 세부적인 내용에 대해서는 아주 잘 정리된 다른 DEVIEW TALK들에서 😊



## 3.1 추천시스템의 꾸준한 고민

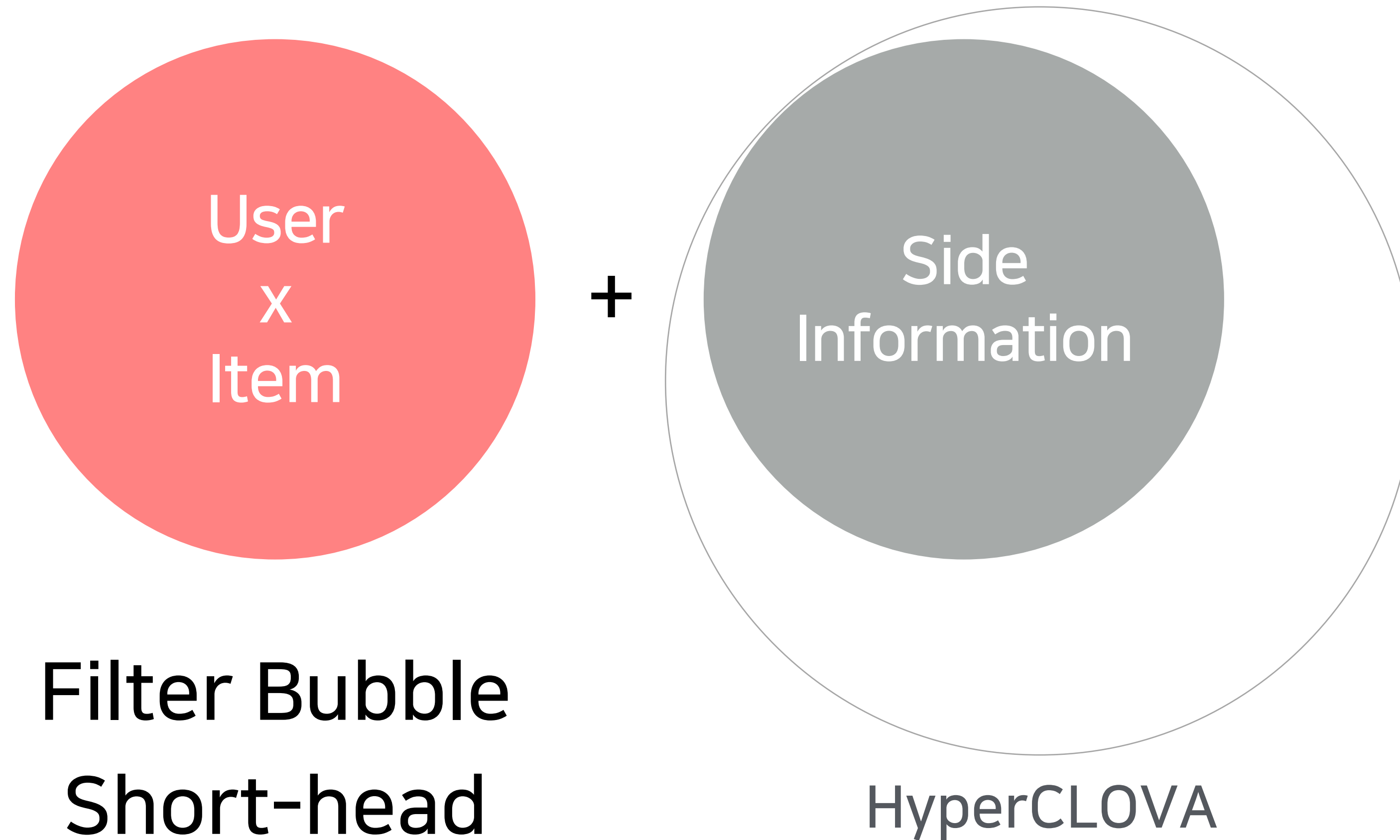
더 다양한 상품을 잘 찾아서  
보여주고 싶다

사용자의 다양한 행동들을  
“잘” 표현해주고 싶다

Diversity,  
Cold-start

Explainable  
AI

# 3.2.1 다양한 상품을 추천하기 위해



카테고리  
브랜드  
상품명  
가격  
...



## 3.2.2 상품명 Summarization

쇼핑 할 때, 상품명이 중요한 요소인 것은 알겠는데...

에어데이즈 마스크 KF94 KF80 KF-AD 1매 새부리형 초소형 소형 중형 대형 조정석

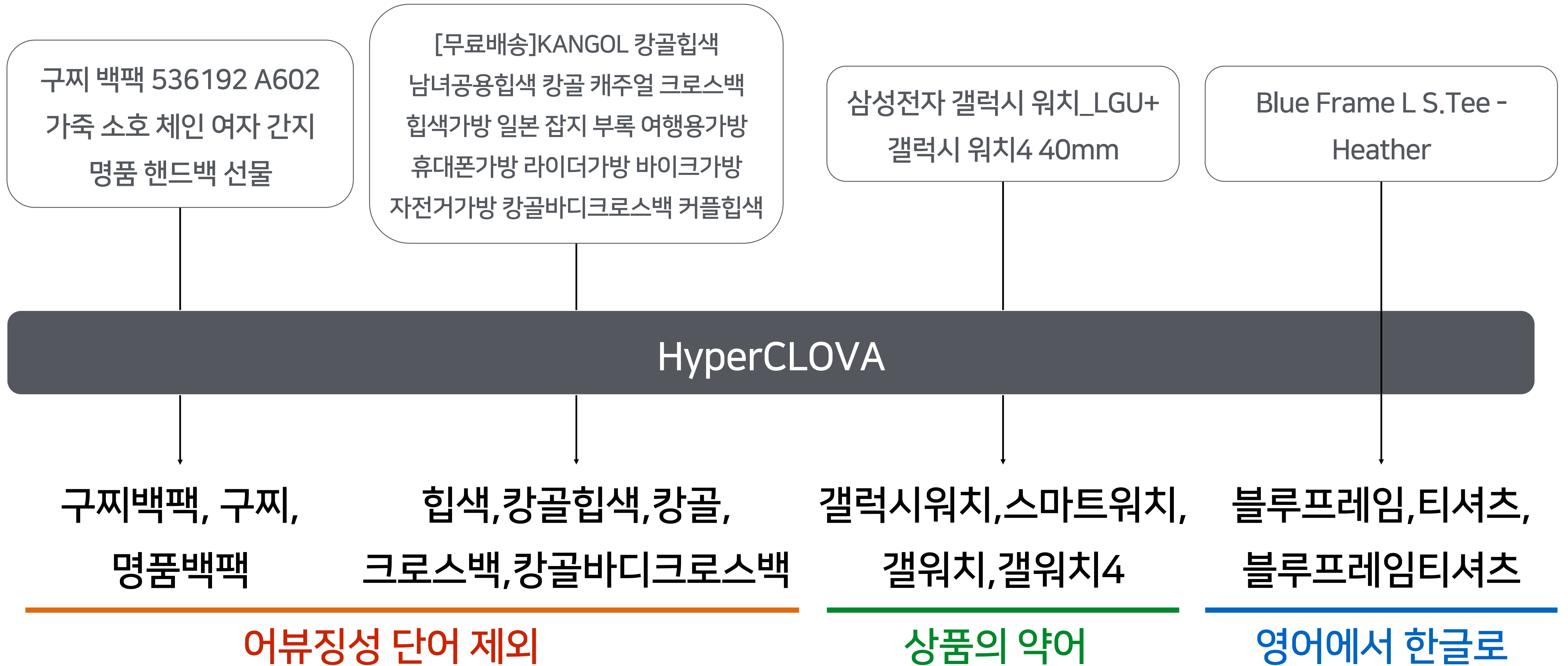
노르딕 led 현대 led 천장 조명 luminaria lamparas de techo 산업 장식 plafon led 거실 침실

5Pcs 생일 파티 케이크 컵케익 Toppers 생일 케이크 추천 케이크 장식품

너무 긴 문장형, 중요 단어 알기 어려움, 어부징성 단어 등

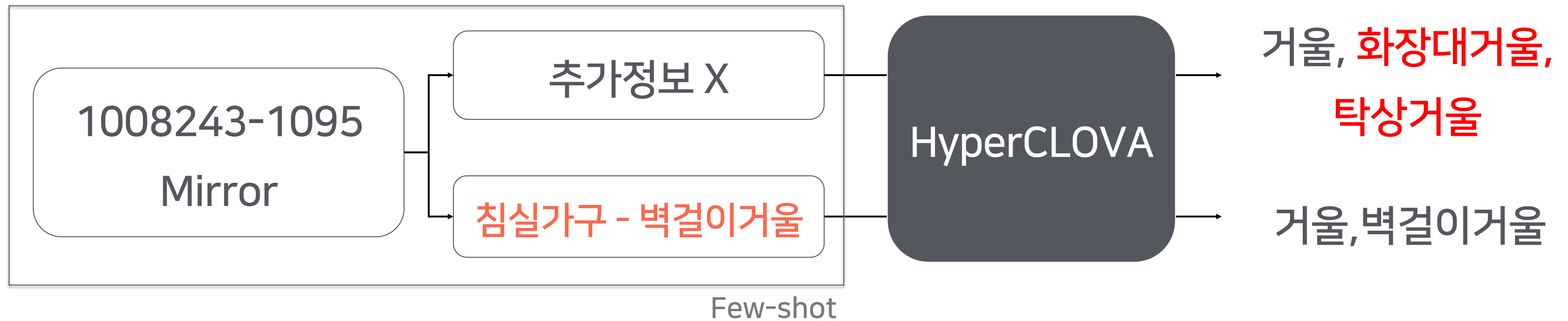
= 좀 더 간결하게 단어 형태로 해당 상품을 표현해 주길 바람

# 3.2.2 상품명 Summarization



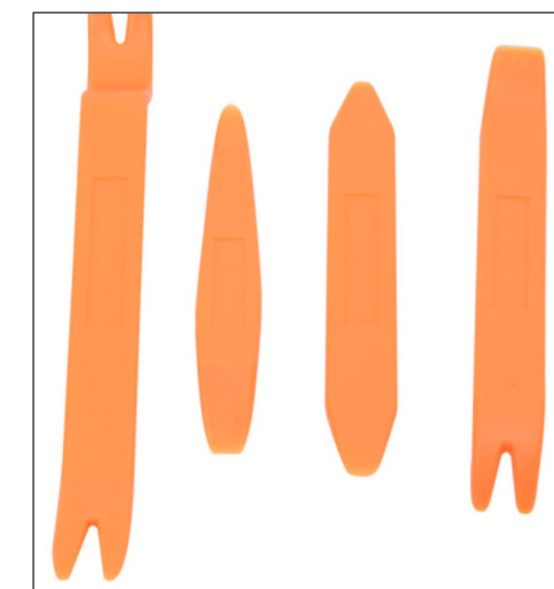
# 3.2.3 상품명 Summarization - 이슈

Few-shot을 적절히 잘 정해 줘야하는...



제목에서 어떠한 정보를 알기 어렵거나,  
카테고리 별로 단어가 중의적인 표현이 있는 경우  
헛갈려 할 때 바로잡아줄 수 있음!

헤라



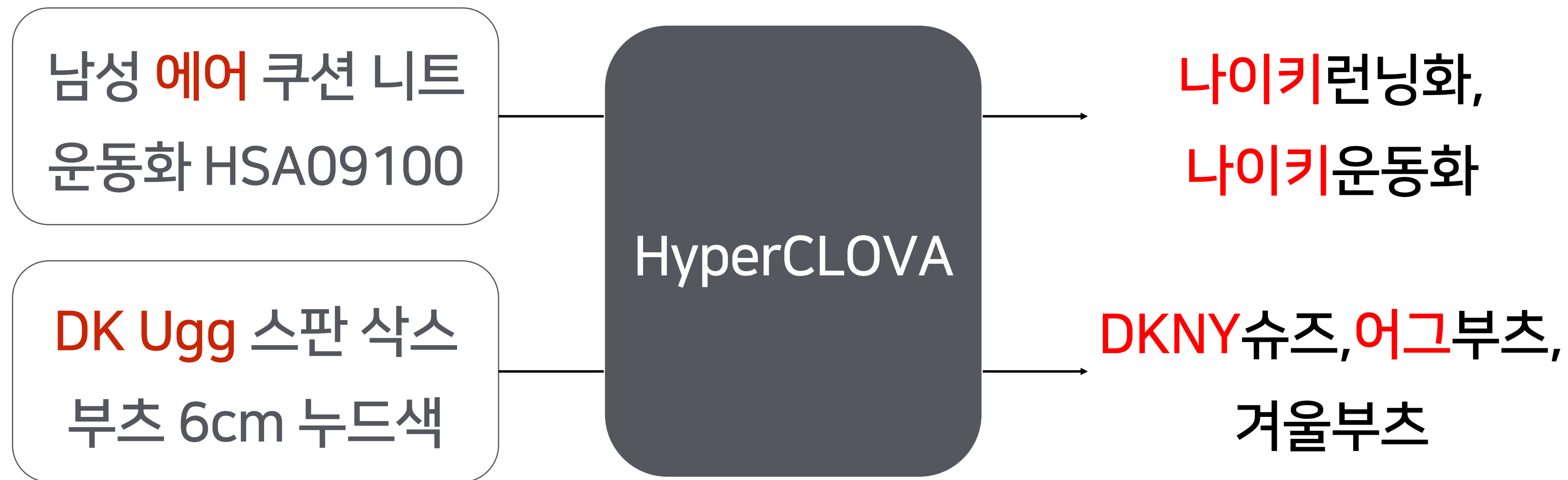
자동차



화장품

## 3.2.3 상품명 Summarization - 이슈

### 확증편향적인 결과



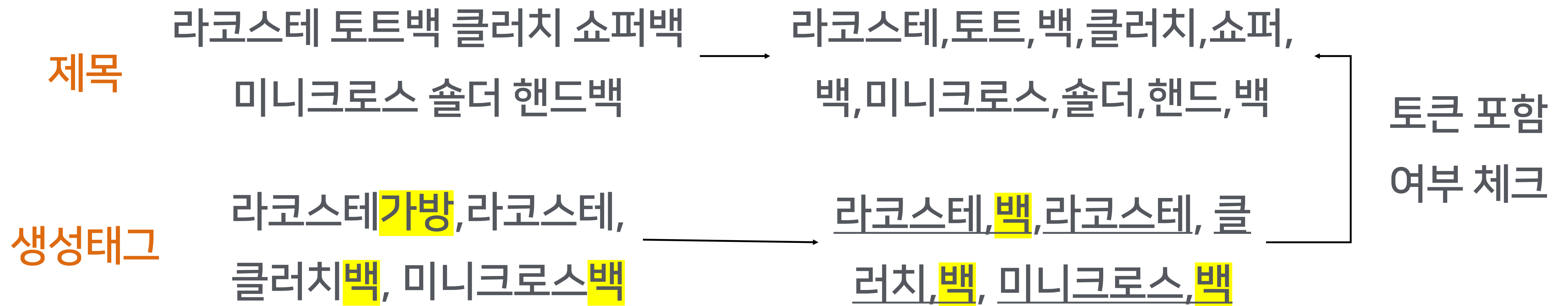
결국 이러한 문제들을 정량적인 수치로 판단할 수 있어야함!

# 3.2.4 상품명 Summarization - 평가

## 데이터 품질 검증 방법 1

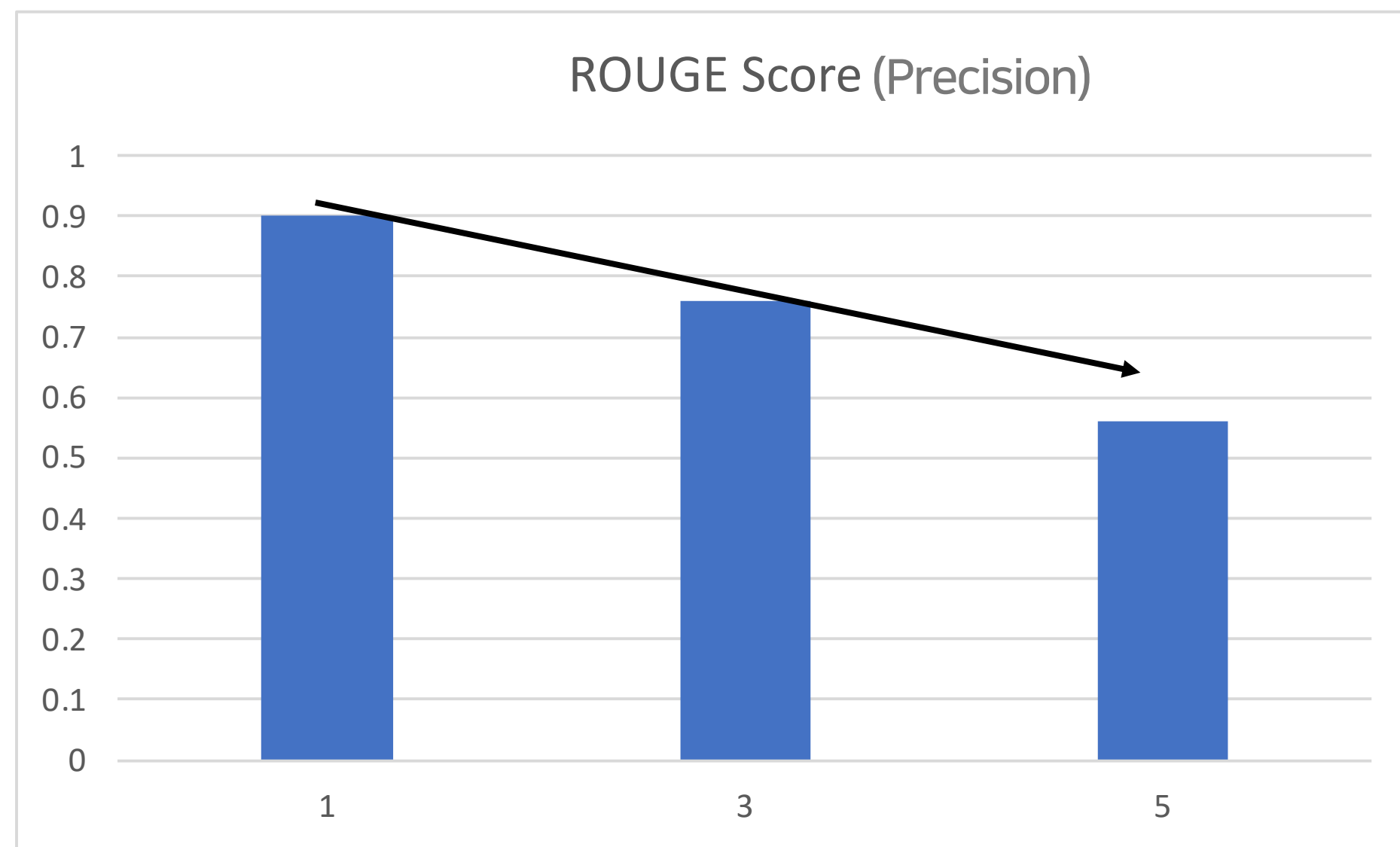
- 입력으로 넣은 데이터 외에는 생성하지 않았는지 체크 (ROUGE)
- 토큰나이징, 영어-한글 변환, 백-가방 같은 유의어 토큰 처리

$$Score_{precision} = \frac{\#제목포함태그}{\#전체태그}$$



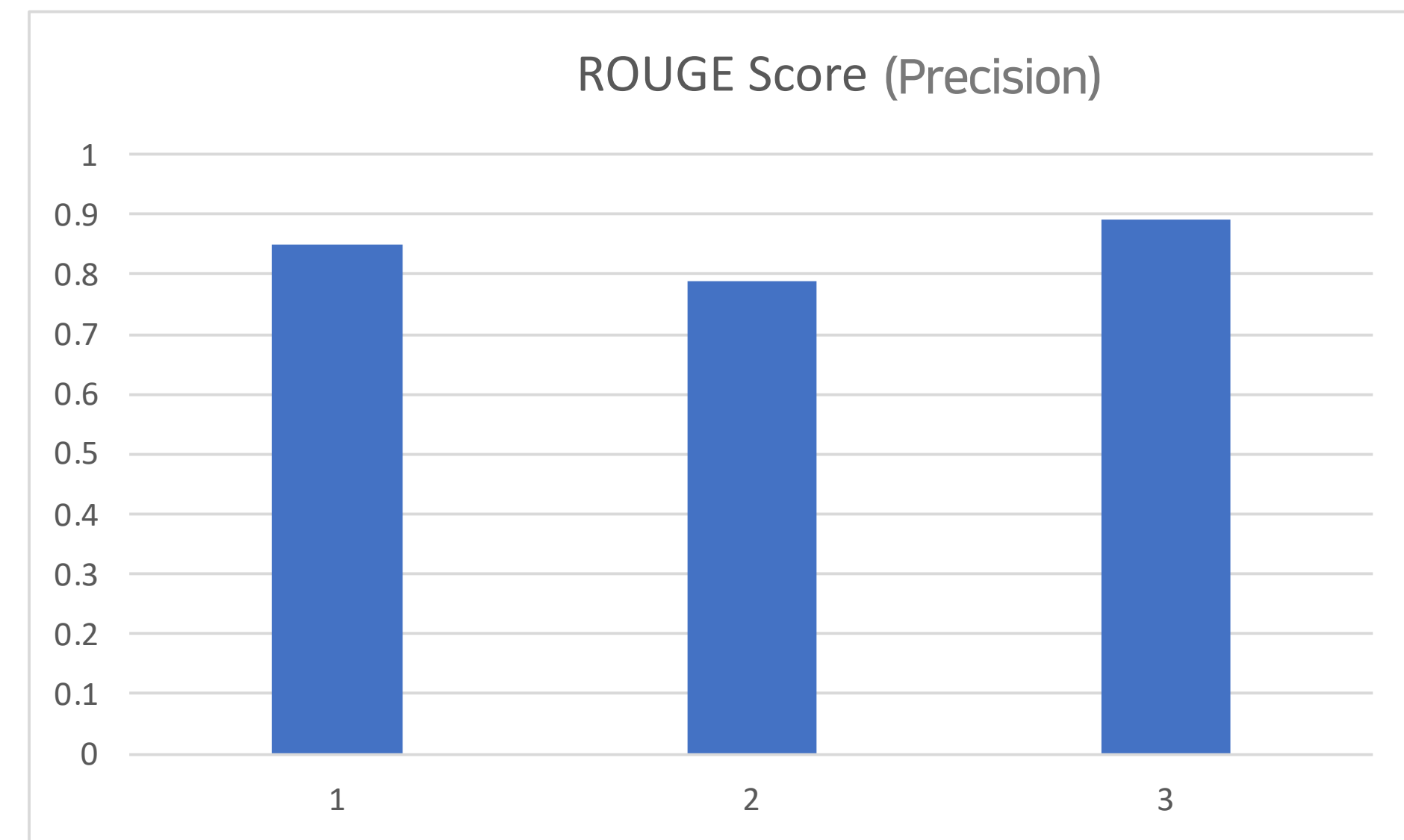
# 3.2.4 상품명 Summarization - 평가

파라미터들에 따라서 평가 값이 많이 바뀌기 때문에 항상 튜닝 필요



Repeat Penalty

(높일수록 제목에 없는 단어들을 마구 생성...)



Few-shot 종류에 따라서



# 3.2.4 상품명 Summarization - 평가

## 데이터 품질 검증 방법 2

- 실제 해당 상품을 검색해서 클릭했을 때 사용한 "검색어"와 직접 비교
- 토큰나이징, 영어-한글 변환, 백-가방 같은 유의어 토큰 처리는 동일

제목

네이처하이크 쓰리동키 클라우드 데스크 M 트윈타워 타프 쉘터 대형쉘터 리빙쉘 텐트 스크린 어닝

생성태그

네이처하이크,네이처하이크데스크,쉘프

실제 검색어

에어쉘터, 거실텐트, 네이처하이크데스크, 쉘프,  
네이처하이크쉘터,네이처하이크대형쉘터

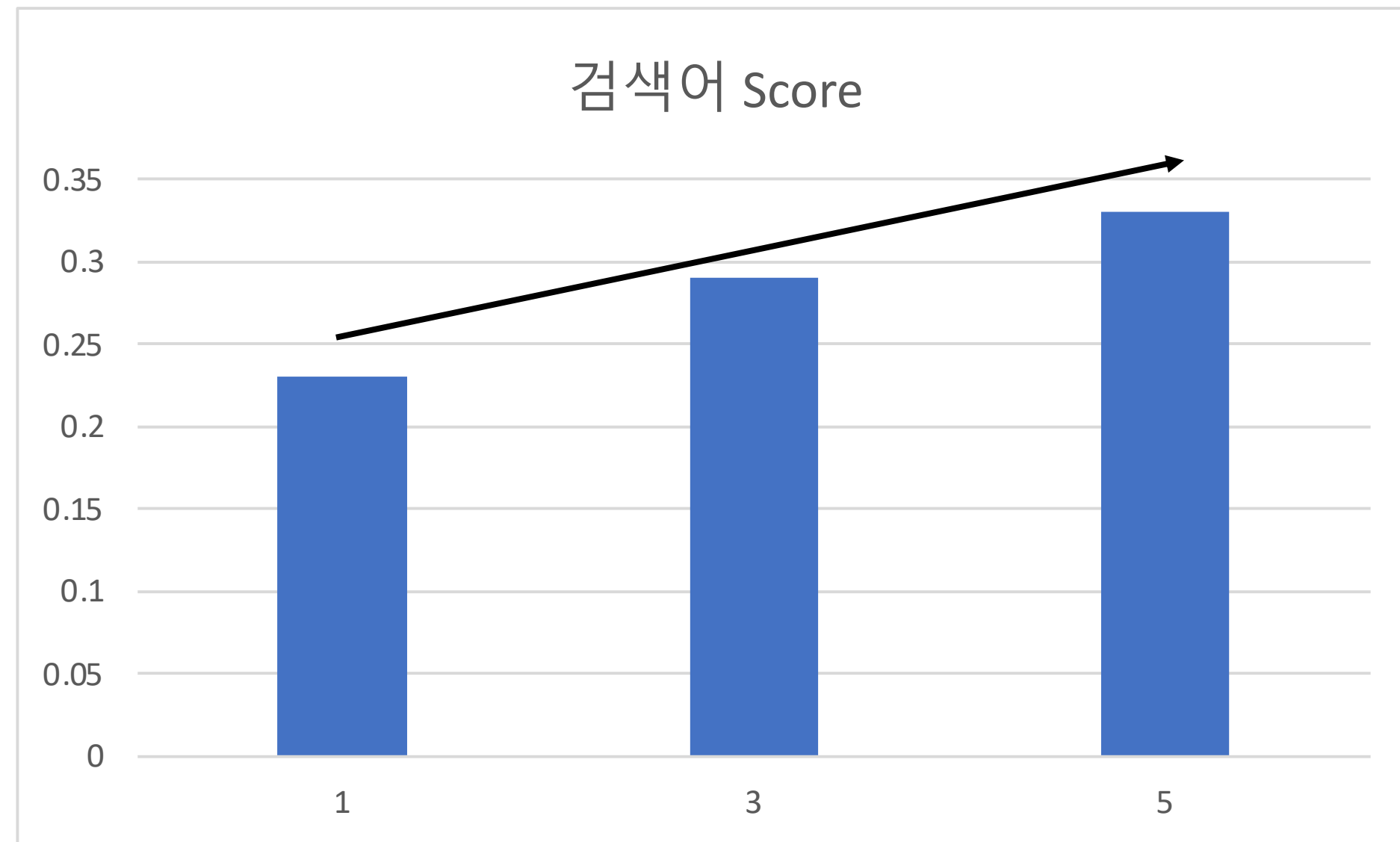


토큰 포함  
여부 체크

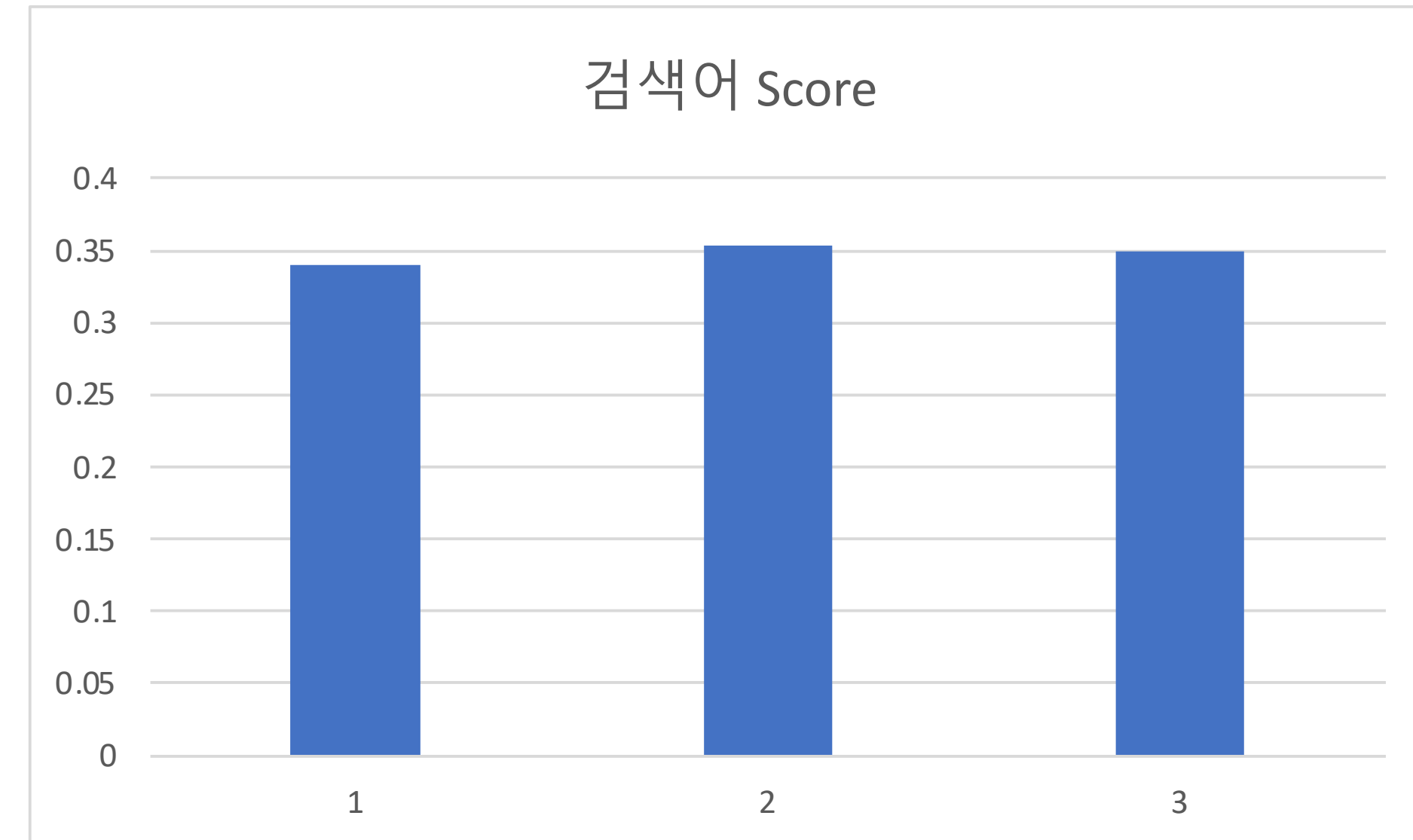
# 3.2.4 상품명 Summarization - 평가

TASK와 아주 Fit하진 않기 때문에 보조적인 지표

$$Score = \frac{\#제목포함검색어}{\#전체검색어}$$



Repeat Penalty



Few-shot 종류에 따라서

(높일수록 제목에 없는 단어들을 마구 생성 하다 보니  
오히려 이 값은 올라가는...)

## 3.2.5 상품명 Summarization - 결론

NLP Task에 상당히 쉽게 접근 가능하고 결과도 좋음! 👍

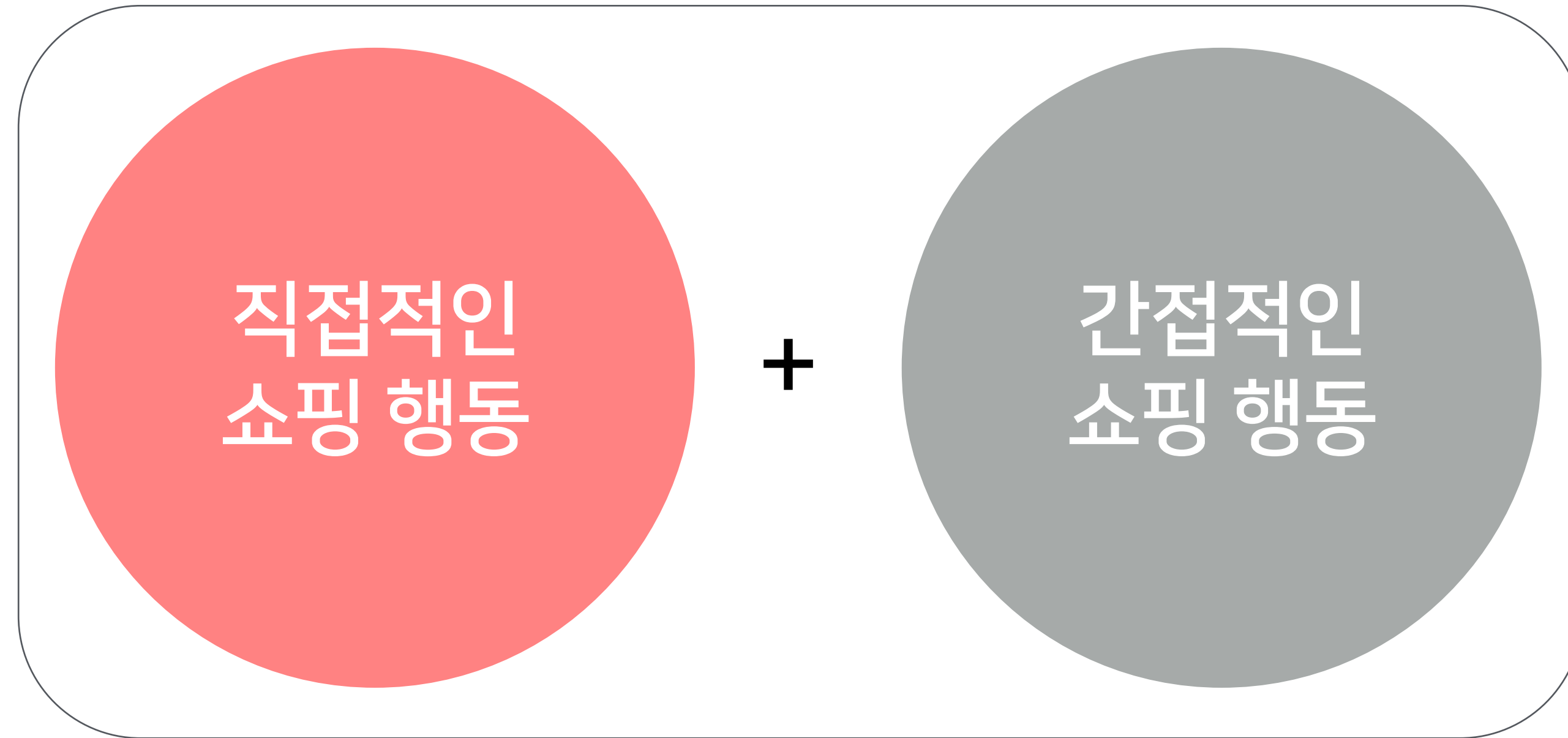
다만, 또 다른 많은 고민거리와 풀어야할 새로운 문제들이 즐비... 🤯

그리고 항상 정답인 평가 방법은 없고, 상황에 따른 평가 지표 필요

상품명 외에도 상품 상세정보 등 훨씬 더 많은 정보들을 가공 시도 중

# 3.3.1 사용자의 다양한 행동들 "잘" 표현

상품 검색  
 상품 클릭  
 찜 하기  
 장바구니  
 구매  
 ...



뉴스  
 블로그  
 카페  
 ...

HyperCLOVA

이런 데이터를 모두 추천 모델링에 활용할 수는 있으나,  
 “지금 #다이어트에 관심있으신가요?” 라고 explicit하게 표현하는건 또 다른 문제...

## 3.3.2 쇼핑 관심사 Generation

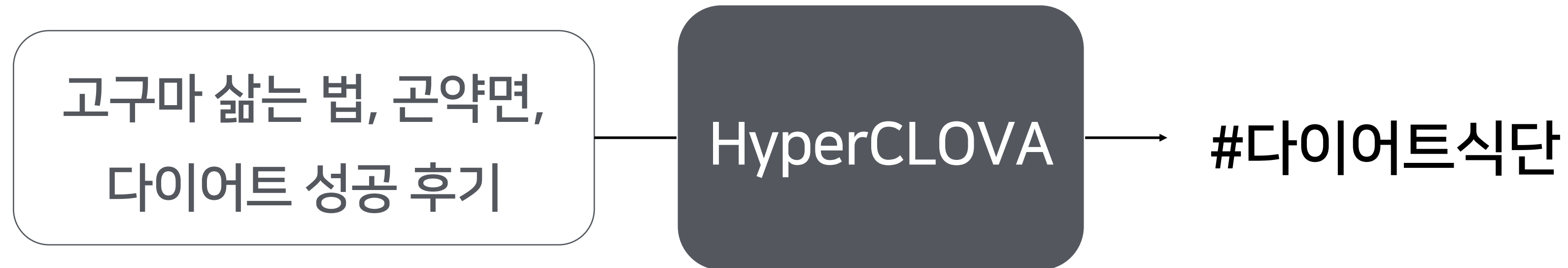
사용자의 여러 검색 기록으로부터 현재 **쇼핑 관심사**를 찾아 보여줌

- 고구마 삶는 법, 곤약면, 다이어트 성공 후기, ... -> #다이어트식단
- 분유포트, 치발기, 6개월아기놀아주는법, ... -> #육아템

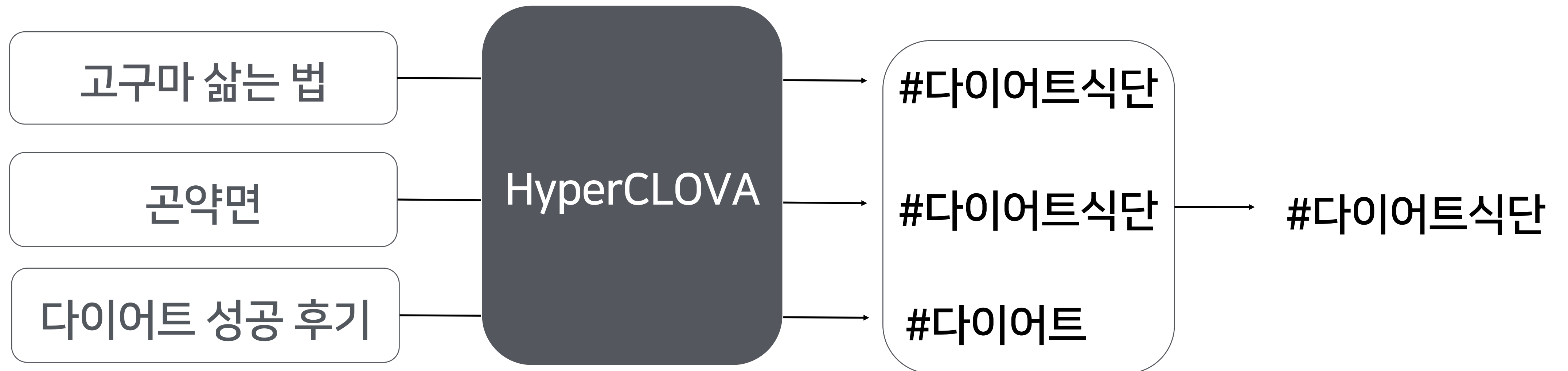
그리고 과거 기록과 현재 쇼핑 관심사를 기반으로 **근거 있는** 상품 추천

- 곤약면과 같이 볼만한 #다이어트식단 제품들: 곤약밥, 두부면, 미역국수, ...

# 3.3.2 쇼핑 관심사 Generation



실 서비스에는 적용이 불가능한 계산량 😱



# 3.3.3 쇼핑 관심사 Generation - 결과

검색 기록 + 상품클릭 등의 활동

**N** 두부면      **N** 비건 베이커리

**N** 미역국수



루리언니님을 위한  
**오늘의 쇼핑 큐레이션** by AiTEMS

나의 관심 주제      테마설정



**다이어트** NEW  
#저탄수빵 #닭가슴살 #미역국수



**피부관리**  
#저탄수빵 #닭가슴살 #미역국수

이 데이터를 활용한 서비스가  
조만간 오픈 예정...! 🤔

9:41      📶      🔋

←      ✕


**#다이어트**      + 관심주제 추가

관심 있으세요?


“두부면”, “미역국수”와 관련있는 쇼핑 태그를 추천드려요

👤 33,545명이 관심있어요      21.8.3 기준

전체      곤약면      소고기      생치즈      키토빵



**99,900원 13%**  
금양식방 비건키토빵 키토빵  
3개 (참깨, 피칸, 흑임자)  
★ 4.9/5 | 리뷰 137



**99,900원 13%**  
금양식방 비건키토빵 키토빵  
3개 (참깨, 피칸, 흑임자)  
★ 4.9/5 | 리뷰 137

## 3.3.4 쇼핑 관심사 Generation - 이슈

### 앞서 상품명 Summarization 에서 생겼던 문제점

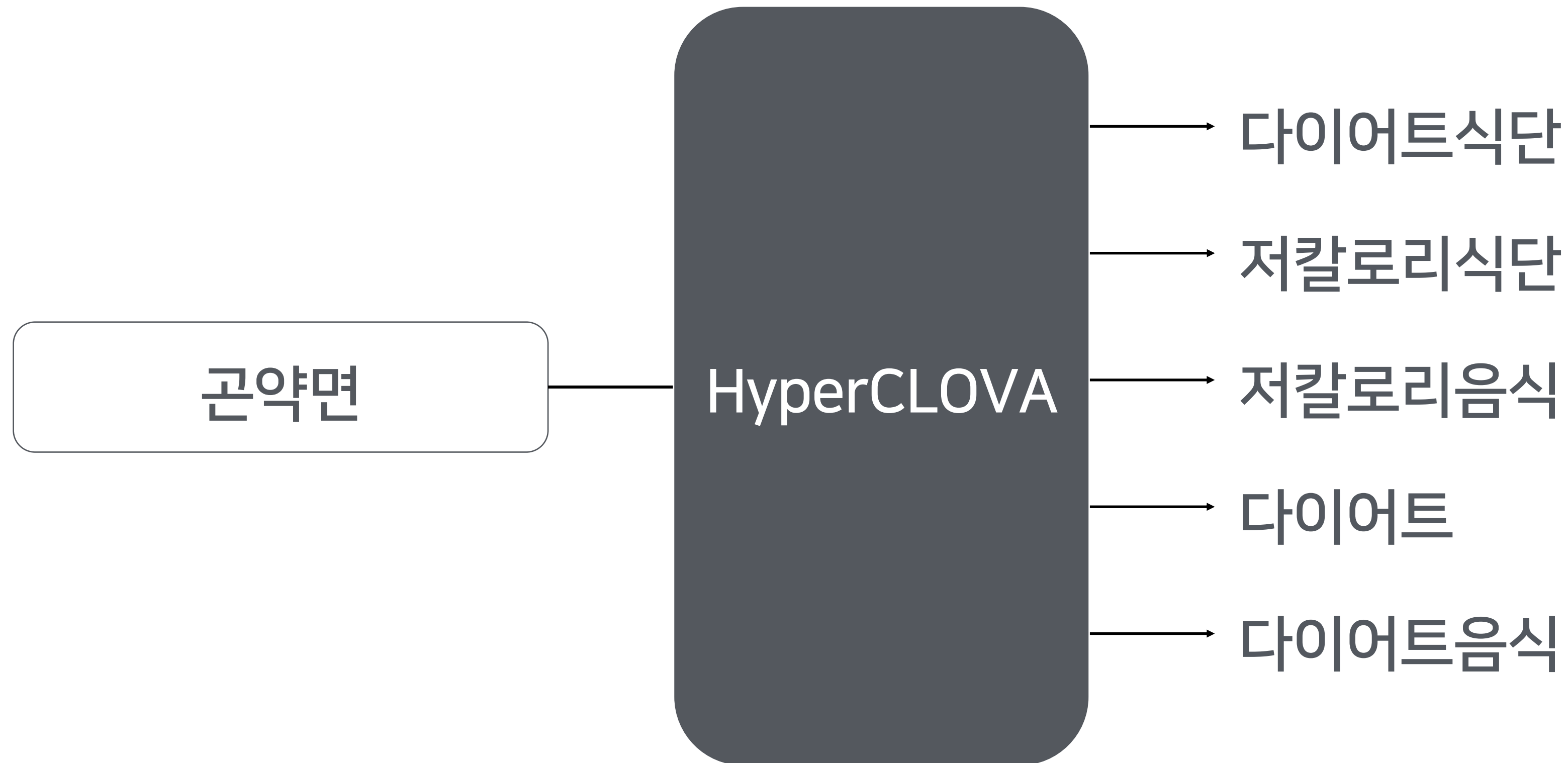
- Few-shot에 영향을 많이 받음 -> 똑같음
- 확장편향적인 문제 -> 의외로 이 Task에서는 문제가 아님!

하지만 또 다른 하나의 문제점이 나오는데...



# 3.3.4 쇼핑 관심사 Generation - 이슈 1

## 생성 모델이 가져온 결과의 다양성 문제



다 맞긴 한데...  
뭔가 좀 더 범주가 큰  
단어도 있고...

워딩은 다르지만 같은  
의미인 단어도 있고...



## 3.3.4 쇼핑 관심사 Generation - 이슈 1

앞선 예시처럼 그나마 깔끔하면 문제가 쉬워지지만 그렇지 않다.

“틀린” 결과를 받는 경우는 없지만, “너무 디테일한” 관심사를 만들어내는 경우가 많음

다이어트식단

샐러드토폭

저칼로리식단

샐러드 드레싱

저칼로리음식

뿐만 아니라...

디톡스

다이어트

건강차

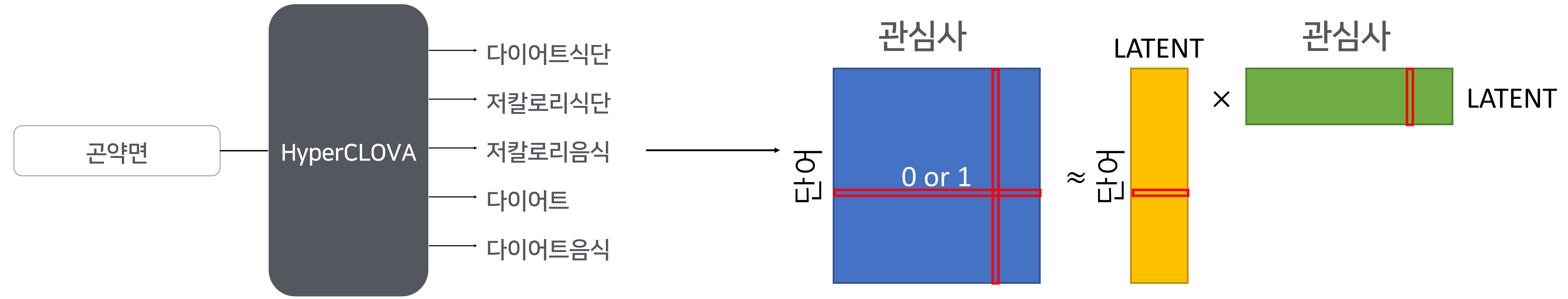
다이어트음식

등등...

또 전부 손으로 하기에는  
너무나 많은 양...

# 3.3.4 쇼핑 관심사 Generation - 해결방법 1

## As Matrix Factorization Problem



단어 별 여러 개의 관심사가 나올 때, 유사 관심사 구하기

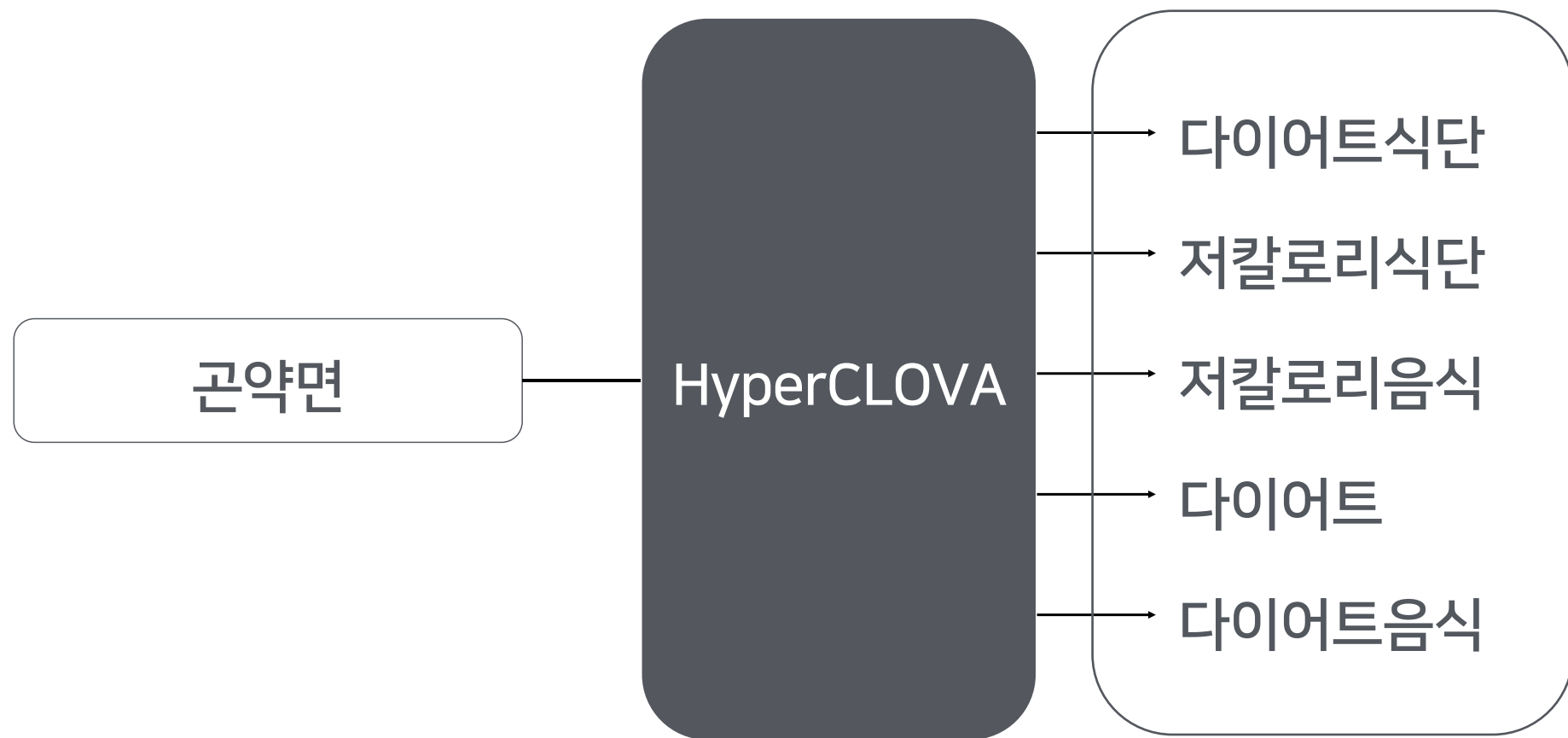
생각보다 너무 Sparse하고  
단어별 생성되는 관심사가 한정됨...

유저 별 여러 개의 상품을 클릭 할 때, 유사 상품 구하기

= 잘 안됨

# 3.3.4 쇼핑 관심사 Generation - 해결방법 1

## As Word2Vec Problem - 1



### Conjecture:

같이 생성된 단어끼리는  
비슷한 단어일 것

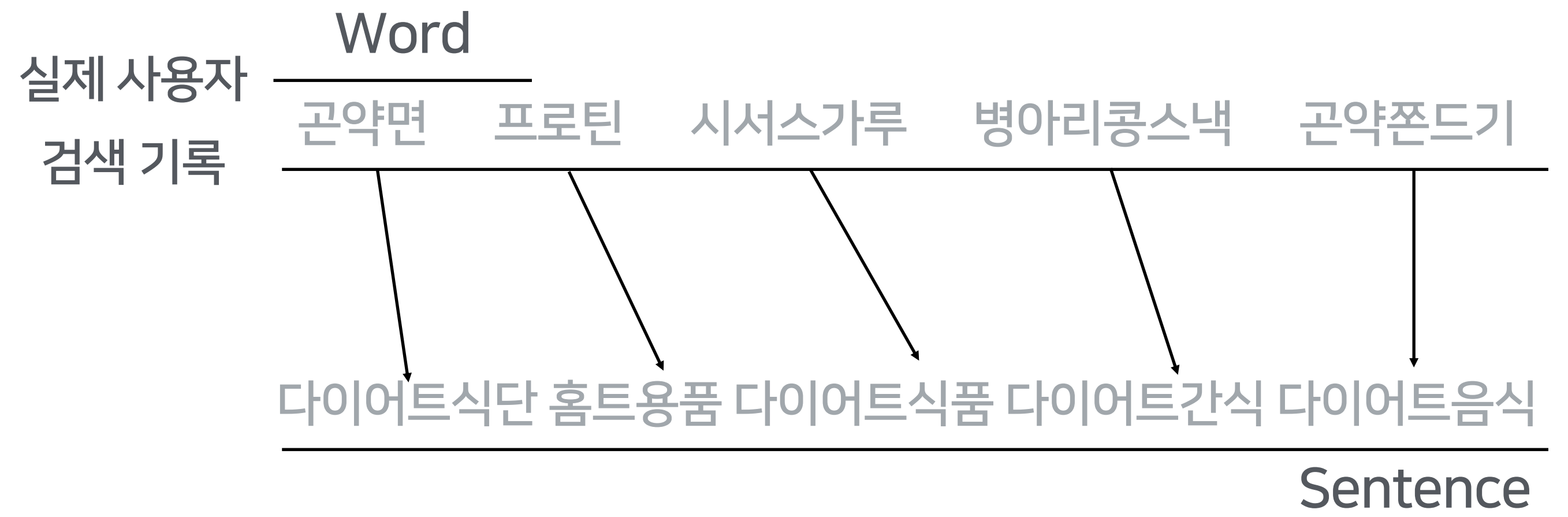
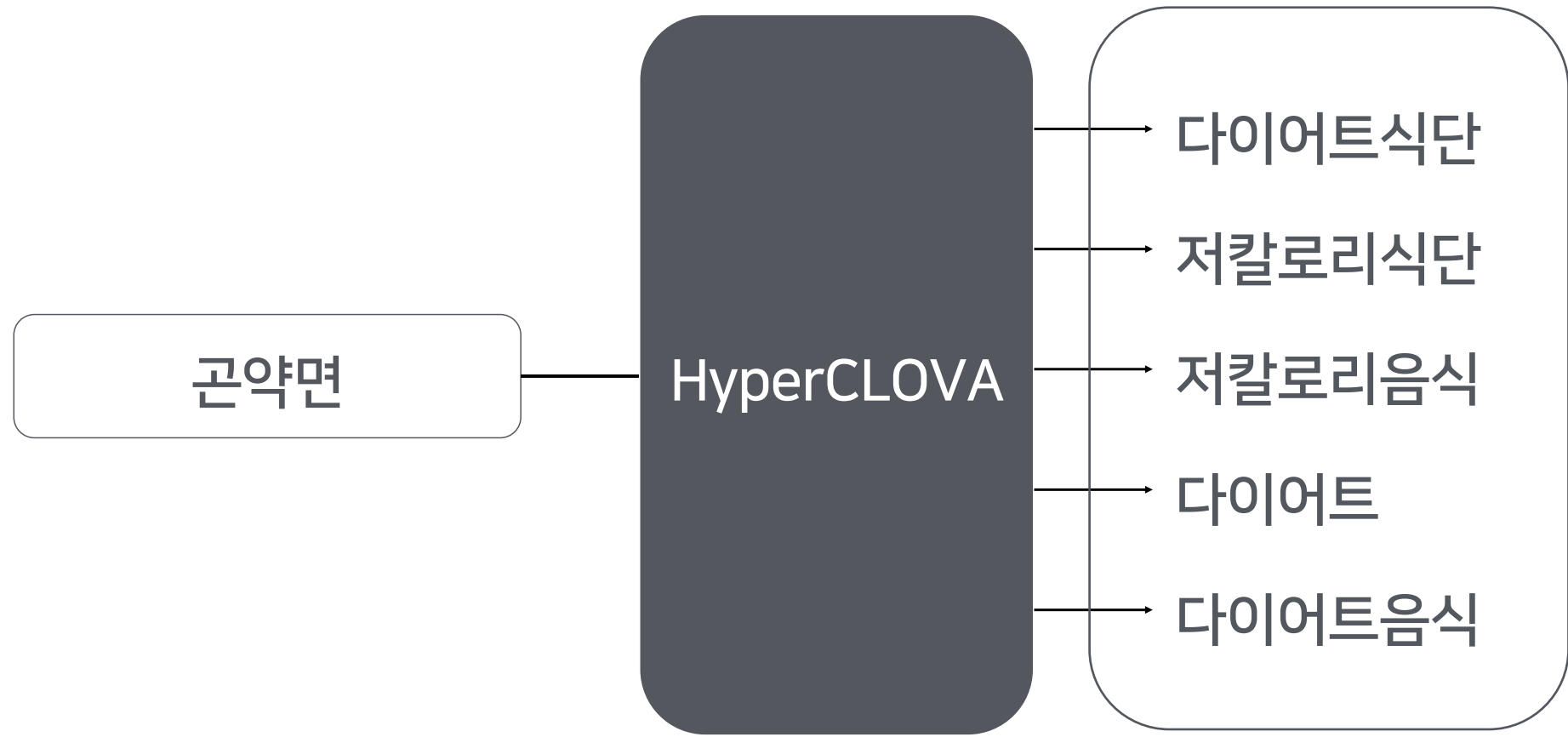
Word
다이어트식단 저칼로리식단 저칼로리음식 다이어트 다이어트음식
Sentence

같이 생성된 관심사들을 모두 Concat하여  
한 문장으로 보고 Word2Vec 학습!

MF보다는 정성적인 결과가 낫지만,  
단어별 풀이 정해져있다면 이상한 결과들이 보임

# 3.3.4 쇼핑 관심사 Generation - 해결방법 1

## As Word2Vec Problem - 2



Conjecture:  
 사람들이 많이 "같이" 검색한  
 결과를 사용하자

사용자가 검색한 기록을 관심사로 치환한 뒤,  
 Word2Vec 학습을 진행!



# 3.3.4 쇼핑 관심사 Generation - 이슈 2

## As Word2Vec Problem - 2 - 결과 샘플

### “다이어트”

- 다이어트음식
- 다이어트음료
- 닭가슴살브랜드
- 체지방관리
- 인바디
- 디톡스다이어트
- 스테비아토마토
- 사과식초
- 다이어트보조제
- 내장지방감소

### “셀프인테리어”

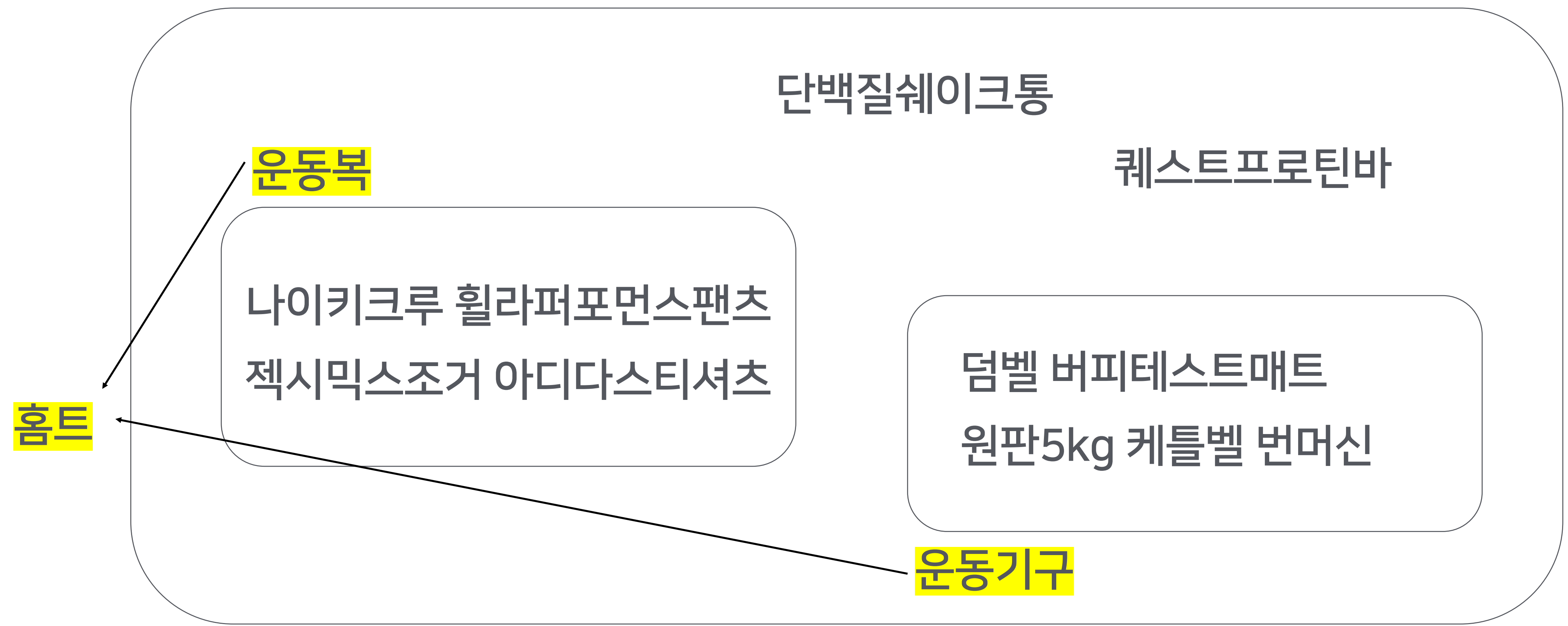
- 조명인테리어
- 방인테리어
- 실내장식
- 천장몰딩
- LED조명
- 생활잡화
- 바닥재
- 페인팅도구
- 인테리어장식
- 리모델링

같이 생성되진 않는 관심사들 이지만,  
실제로 같이 많이 봄직한 유사 관심사 품질

근데 이렇게 디테일한 관심사들이 생성되는데,  
“다이어트”, “셀프인테리어”는 어떻게 뽑나?

# 3.3.4 쇼핑 관심사 Generation - 해결방법 2

관심사에 속한 단어들 간의 포함관계로부터 관계 파악



# 3.3.4 쇼핑 관심사 Generation - 해결방법 2

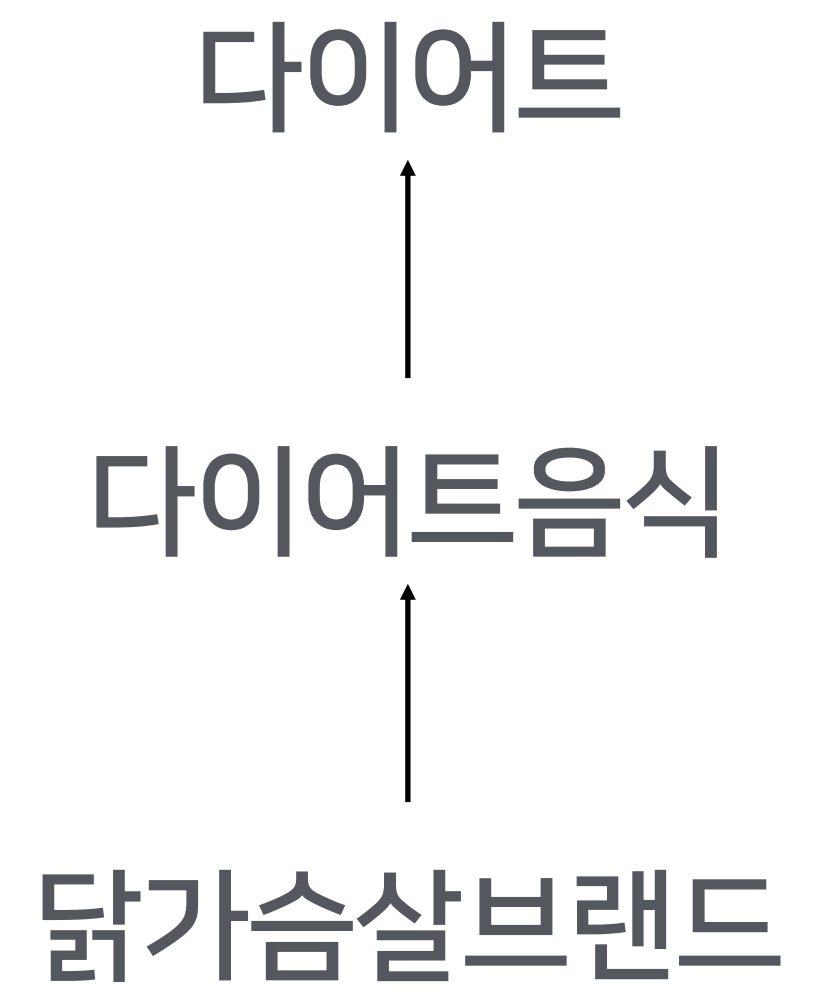
그렇게 Word2Vec 결과와 합쳐서, 일정 거리 내 포함 관계를 모두 기록

## “다이어트”

- 다이어트음식
- 다이어트음료
- 닭가슴살브랜드
- 체지방관리
- 인바디
- 디톡스다이어트
- 스테비아토마토
- 사과식초
- 다이어트보조제
- 내장지방감소

## “다이어트음식”

- 다이어트
- 다이어트식단
- 식단관리
- 헬스식단
- 간헐적단식
- 닭가슴살브랜드
- 스테비아토마토
- 매운맛소스
- 체중조절
- 체지방측정기





## 3.3.4 쇼핑 관심사 Generation - 해결방법 2

하지만 이것만으로는 100% 자동화는 어려웠음

- 너무 넓은 범주의 관심사가 생김. Ex) 데일리룩, 데일리메이크업, 튜닝
- 항상 같은 Level의 관심사만이 정답은 아님. Ex) 다이어트 vs 와인안주

위 방법들로 많이 정제하되, 최종 결정은 사람이 개입하는 반자동 형태로!

- **수만개** 수준의 데이터에서 **수백개** 수준의 데이터로 범위를 좁혀서,
- 사람이 손쉽게 검토하여 사용할 수 있도록 반자동화.

# 3.3.5 쇼핑 관심사 Generation - 결론

기존에 가지고 있지 않던 데이터를 품질 좋게 생성 가능!

덕분에 추천 분야에서 상당히 어려운 문제인 "Explainability"를 어느정도 해소할 수 있었고,

다만, 역시나 서비스화를 위해서는 튜닝하고 후 처리해야 할 작업이 ... 🤯



## 3.4 상품 추천 문장 개인화 맛보기

더 나아가 유저들의 여러 쇼핑 관심사와 각 사용자 정보를 결합한다면?

“김동현님을 위한 추천상품”



“이 상품 어떠세요?”

“김동현님이 좋아할만한 상품”

“**다이어트**를 위해 **운동 의류**를 보고계신가요?” #다이어트  
#운동의류

“**화장품**이 많은 분들의 필수템! 작지만 알찬 **미니냉장고** 둘러보세요!” #화장품  
#미니냉장고

“작은 사이즈로 **공간 부담없이** 사용할 수 있는 **미니냉장고**” #원룸인테리어  
#미니냉장고



## 3.4 상품 추천 문장 개인화 맛보기

아직까지는 여러 문제점들이 있어서 테스트 중

- 나의 사람 관계에 맞지 않는 문구들: 남친/여친, 아내/남편, 육아맘 등
- 계절에 맞지 않는 문구: "요즘 같은 여름에~"
- 그 외 서비스적으로 부적합한 문장이나 용어 포함
- 학습된 데이터로부터 거짓된 정보를 내포하는 문장을 만드는 경우

# 3.5 AiTEMS X HyperCLOVA

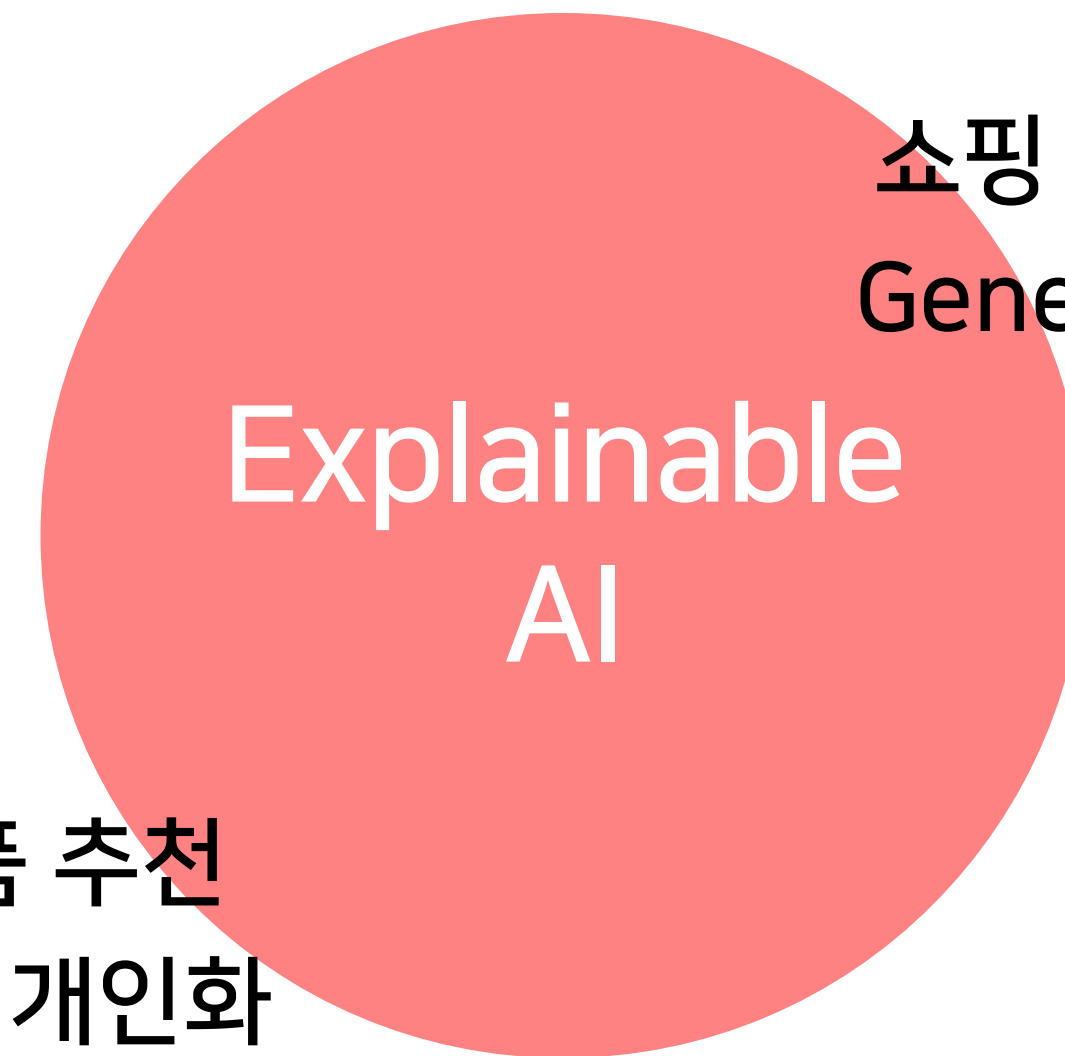
더 다양한 상품을 잘 찾아서  
보여주고 싶다

사용자의 다양한 행동들을  
“잘” 표현해주고 싶다

And more...



상품명  
Summarization



상품 추천  
문장 개인화

쇼핑 관심사  
Generation

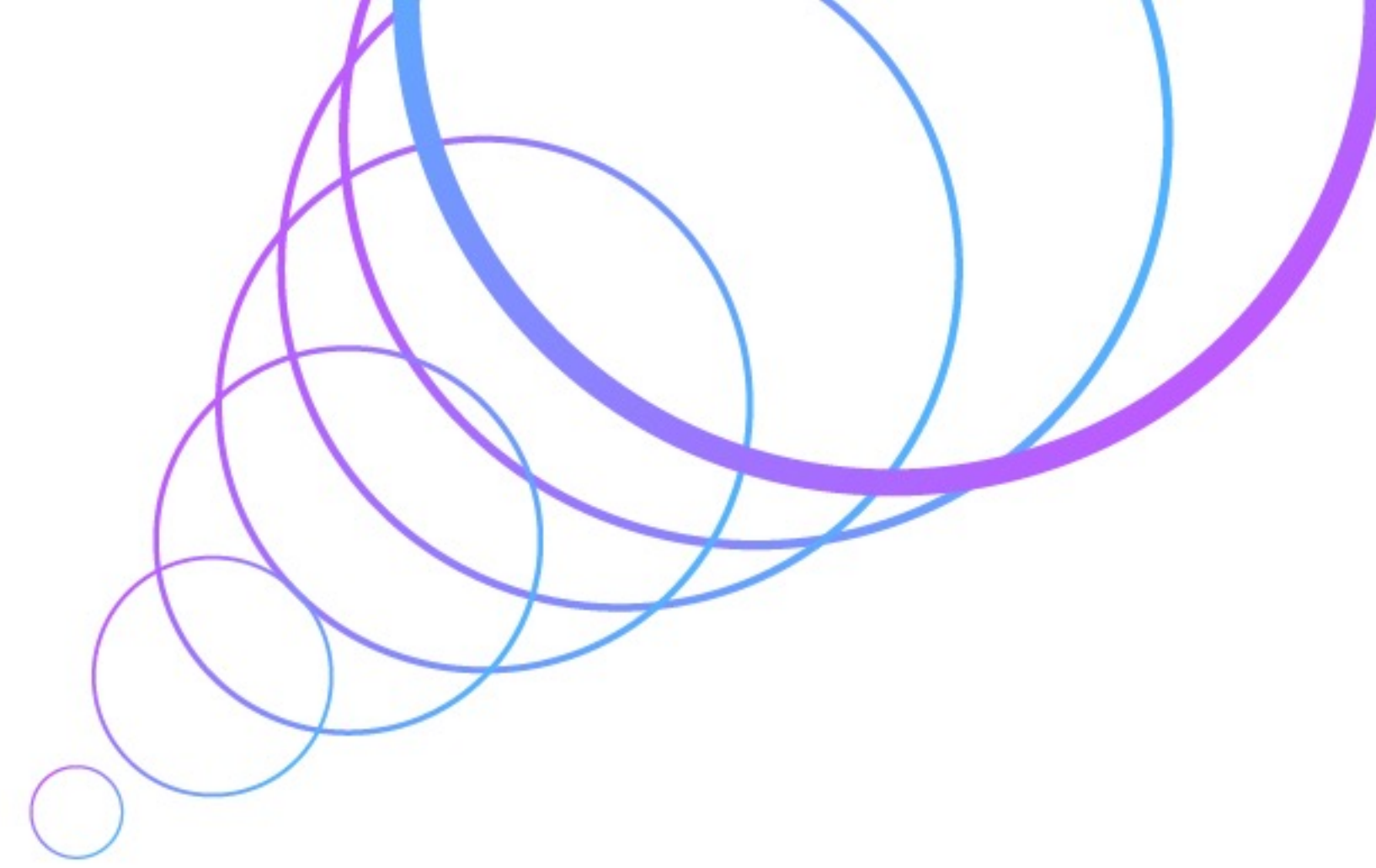
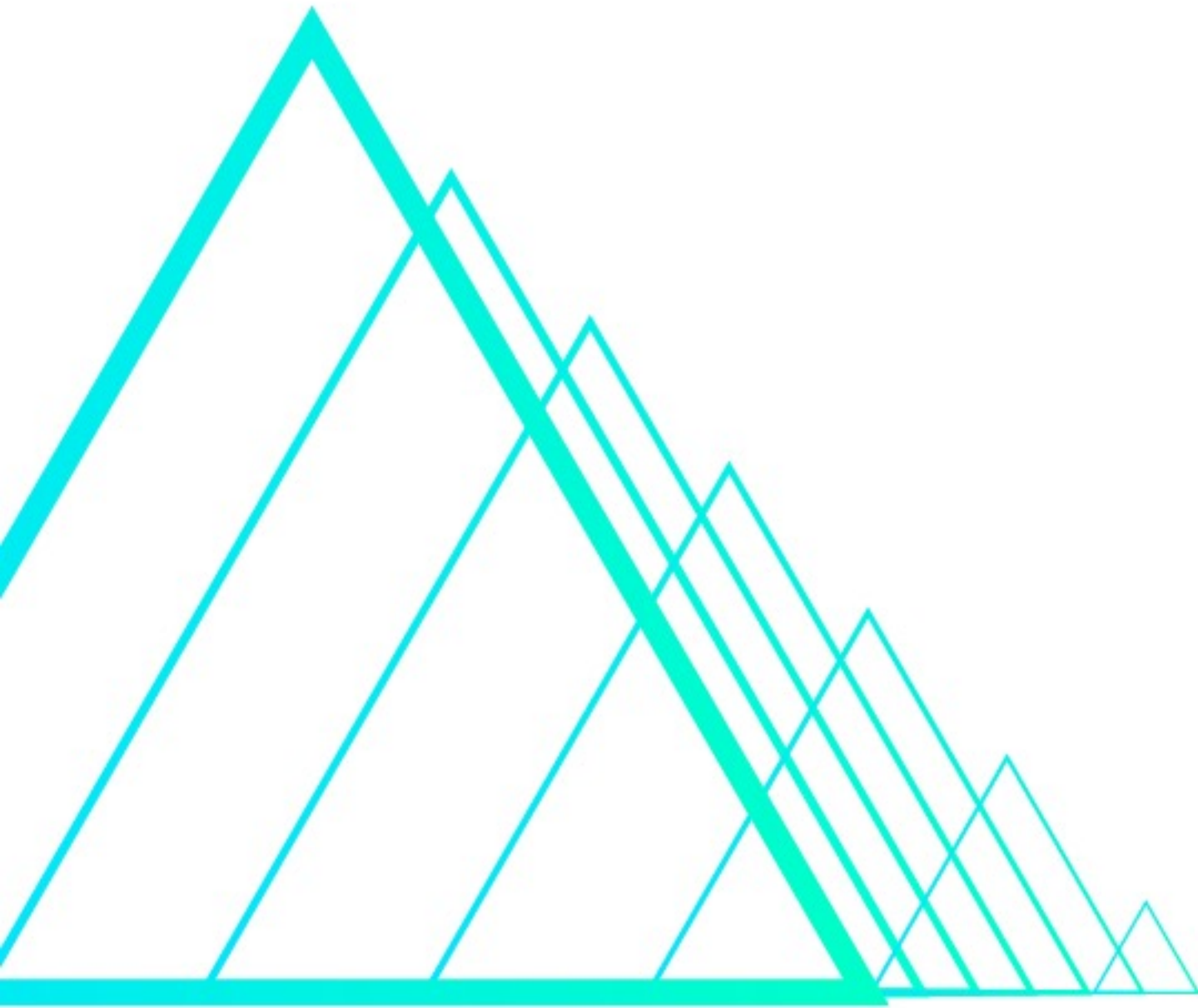
## 4. Future works

결국은 “더욱 개인화된 사용자 경험”을 주기 위한 여러 모델 연구 개발

- 개인화된 추천 문구
- 익숙함과 새로움의 적절한 비율 (Exploitation vs Exploration)
- 더 많은 사용자의 기록들을 어떻게 더 “잘” 사용할 것인지...

그리고 모든 연구/서비스 개발 시에 가장 어려운 부분은 역시나 평가

- 모델과 서비스 모두 커지고 복잡해짐에 따라 평가 방법도 점점 더 어려워지고 있다...



**Thank You**

