

<에이블리>

-블로그, 인스타그램, 유튜브 등 흩어져 있던 개인 셀러들을 한곳에 모은 쇼핑 플랫폼 모바일 앱

2018년 3월에 만들어진 신생앱이지만 불과 1년사이 11배 성장

에이블리가 강조하는 경쟁력은 데이터 기반 추천 서비스

연결을 만드는 '개인화 추천' -> 에이블리는 특정 소비자가 선호하는 컨셉에 맞는 판매자를 잘 연결해줌

좋은 추천 서비스를 만들기 위해서는 1. 좋은 알고리즘 개발 2. 데이터 취합 이 필요함

에이블리는 '상품찜' 데이터를 중요하게 생각하고 왓차에서 영화 평가를 유도해서 특정인이 선호하는 영화 추천을 위한 기반 데이터를 모았듯, (강석훈 대표가 '왓차'라는 영화 추천 서비스의 공동 창업자였고 현재 에이블리 팀은 왓차에서 it를 했던 팀) 에이블리는 상품찜 평가를 유도하여 사용자의 패션 선호 데이터를 모음. 한달에 상품찜 데이터만 1억개 이상이 쌓이고 있다고 함.

또 중요하게 보는 데이터는 상품 자체가 가지고 있는 메타 데이터. 상품 사진과 해당 상품을 묘사하는 텍스트 설명, 어떤 판매자가 판매자가 판매하고 있는지, 특정 판매자의 체형 정보는 어떠한지 수집.

이렇게 고객의 패션 선호 데이터와 상품의 메타 데이터를 연결하면 고객이 어떤 판매자를 좋아하는지, 어떤 고객이 선호하는 사진은 무엇인지 알 수 있음.

에이블리의 추천 서비스의 기반은 아마존 웹 서비스 AWS의 아마존 퍼스널라이즈.

아마존 퍼스널라이즈이란?

아마존닷컴에서 수년간 학습시킨 머신러닝 기술을 고객사 앱에 접목할 수 있도록 지원하는 서비스. 기계 학습과 관련한 경험이 없는 개발자도 자신의 애플리케이션에 개인 맞춤 항목을 생성할 수 있는 완전 관리형 인공지능 서비스. (아마존 퍼스널라이즈 서비스 기능에 대한 상세 동영상 있음)

국내 사례

롯데마트는 Amazon Personalize를 도입해 자체 모바일앱인 'M쿠폰'을 통해 매장 이용고객 대상으로 고차원의 맞춤형 상품 추천 서비스를 제공하고 있습니다. 온-오프라인 구매 데이터를 기반한 기계 학습 결과를 통해 M쿠폰 기반의 1:1 구매 상품 추천을 진행하고 있으며, 보다 정교하고 적중도가 높은 고객 제안을 위해 아마존 퍼스널라이즈를 도입하였습니다. 작년 7월부터 2개월 동안 10만명 이상의 고객 대상으로 1:1 상품 추천 서비스를 시범적으로 운영한 결과, 기존보다 2배 이상 높은 고객 반응율을 이끌어 냈습니다.

롯데마트 빅데이터팀 신재현 팀장은 "이번에 도입한 Amazon Personalize는 아마존닷컴에서 실제로 사용중인 추천 알고리즘"이라며, "광고가 범람하는 시대에 AI를 이용해 개인에게 맞는 최적의 상품을 적절한 시점에 추천 할 수 있게 됐다"고 전했습니다. 정식 출시를 통해 30만 명을 대상으로 상품 추천 서비스가 추천하는 상품을 가격가에 구매할 수 있는 모바일 쿠폰을 M쿠폰앱을 통해 제공하고 있습니다.

인터파크 역시 Amazon Personalize를 통해 특정 옷을 사면 어울리는 패션 액세서리를 추천하거나, 동일 상품 구매 고객이 추가로 구매한 상품을 추천하는 등 활용 가능한 모든 데이터를 복합적으로 조합해 개개인에 최적화된 상품을 제안하는 기능을 구현하고 있습니다.

기존의 과거 구매 내역이나 성별, 연령대 등 일정 카테고리로 고객군을 나눠 단순히 상품을 추천하는 것과는 차별화되는 방식입니다. 클릭, 페이지 뷰, 로그인, 구매, 아이템 재고 등 인터파크 웹사이트의 액티비티 스트림 정보를 기반으로 특징을 식별, 적절한 알고리즘을 선택해 자동으로 개인화 모델을 만들고 있습니다.



[카테고리 없음](#)

쿠팡 추천시스템 분석

A_sparagus | 2020. 3. 26. 16:51

참고자료

[쿠팡 추천 시스템 2년간의 변천사 \(상품추천에서 실시간 개인화로\)](#)

[ppt 자료](#)

과거, 모델 중심의 플랫폼

단일 모델 혹은 복수의 모델이 상품 추천의 모든 역할 수행

모델 복잡 & 서버 단순

모델은 데이터를 입력받아 복잡한 알고리즘을 통해 학습되며, 이 학습 결과를 통해 새로운 input의 결과를 예측
서버는 모델이 내놓은 결과물을 정렬하는 등의 역할만 수행하면 됨.

모델 변경이 필요하거나, 데이터에 문제가 있거나, 학습을 하루 잘못 시키거나, 새로운 필터를 적용시킬 필요가 있거나(ex.성인 용품) 하면 모델을 다시 학습시켜야 한다.

그렇기 때문에 다음과 같은 문제가 발생한다.

- 1) 모델 변경에 따라 길어지는 파이프라인(필터, 부스팅, ...)
- 2) 추가 요청사항에 대한 유연성이 떨어진다.
- 3) 완성 전까지 결과를 알 수 없다.
- 4) 개발에 시간이 오래 걸린다. (새로운 추천 서비스 개발에 한 달 정도 걸렸다고 함)
- 5) 상품, 유저 정보 중 일부는 시스템 서빙 시에만 접근이 가능하기 때문에, 추천 시스템을 개발하려고 할 때 접근이 힘들다.
- 6) 점진적 개선이 힘들다.(모델의 재활용이 어렵다.) 지난 모델은 폐기하기 때문

현재, 서비스와 모델을 분리하는 플랫폼

서비스가 모델로부터 독립. 모델은 Feature만 추출하고, 서비스가 그 결과들을 기반으로 상품 추천 수행

모델 단순 & 서버 복잡

모델은 Feature를 추출하는 역할만 할 만큼 단순. 대신 서버 측에서 파라미터 튜닝 모델을 학습시키는 등 일을 많이 함

이렇게 할 경우 다음과 같은 장점이 있다.


- 1) 빠른 서비스 개발. 피처와 가중치 등의 조합에 따라 새로운 서비스 개발이 가능하다.
- 2) 만들어 놓은 Feature는 다른 서비스에서도 사용 가능
- 3) 서비스에 문제가 생기더라도 쿼리를 튜닝해서 문제를 해결할 수 있다.

추천의 예시

1. 유저 기반 개인화

오성민님을 위한 추천상품

더보기 >




블루다이아몬드 아몬드 브리즈 언스위트, 190ml, 24개

13,140원

로켓배송

추가 할인쿠폰




구달 청귤 비타C 잡티세럼 기획 세트, 세럼 30 ml + ...

13,270원

로켓와우

추가 할인쿠폰



바닐라코 뉴 클린 잇 렌징 밤 오리지널 대

15,780원

로켓와우

추가 할인쿠폰

2. 개인화 할인상품

유저 기반 개인화에서 필터와 부스팅을 활용해 나만을 위한 특가 상품 추천 등 가능

특가일 때 구매하세요!

더보기 >



마블 큐티 보조배터리 5000 + 8핀젠더, SPE-PBC50...

즉시할인가 3% 14,900원

14,450원

로켓와우



라인프렌즈 보조배터리 5000mAh, TEM-P5-5...

즉시할인가 29% 26,000원

17,650원

로켓와우



비티21 원형케이블 C, VAN, 1개

즉시할인가 30% 12,130원


8,460원

로켓와우

3. 시간 기반 인기상품(카테고리 추천)

요즘 인기있는 방향제/디퓨저


더보기 >



릴리브디퓨저 비행기조종사
차량용 방향제 퓨어클린 브...

29,000원


로켓와우



VIP 하이폴리머 차량용 방향
제 팬던트형 바질향 No6,...

14,800원

로켓배송



산도깨비 실내 비치향
향 공기 방향제 자동


8,000원

4. 인기 있는 PB

상품 attribute에 필터를 건다면 PB상품 추천도 가능

쿠팡 only 상품

더보기 >



쿠팡 브랜드 - 코멧 순백 3겹
라벤더 바닐라 롤화장지 3...

9,490원

로켓배송



쿠팡 브랜드 - 탐사 스파클링
플레인 탄산수, 500ml, 2...

8,490원

로켓배송



코멧 순수한 오리지널
물티슈 캡형, 100매

9,690원

로켓배송


추가 할인쿠폰

5. 카트 분석

카트에 담긴 모든 아이템을 고려해 아이템 추천

함께 사는 상품을 추천하거나, 무료 배송을 마치기 위해 더하면 좋은 상품들을 추천해주기도 함.

함께 구매하시면 좋을 것 같아요




쿠팡 브랜드 - 코멧 홈 화장
실 규조토 발매트, 그레이

8,990원

로켓와우

★★★★★ (2,891)




코튼클라우드 호텔수건
200g 40수 코마사, 아이...

17,650원

로켓와우

★★★★★ (2,424)



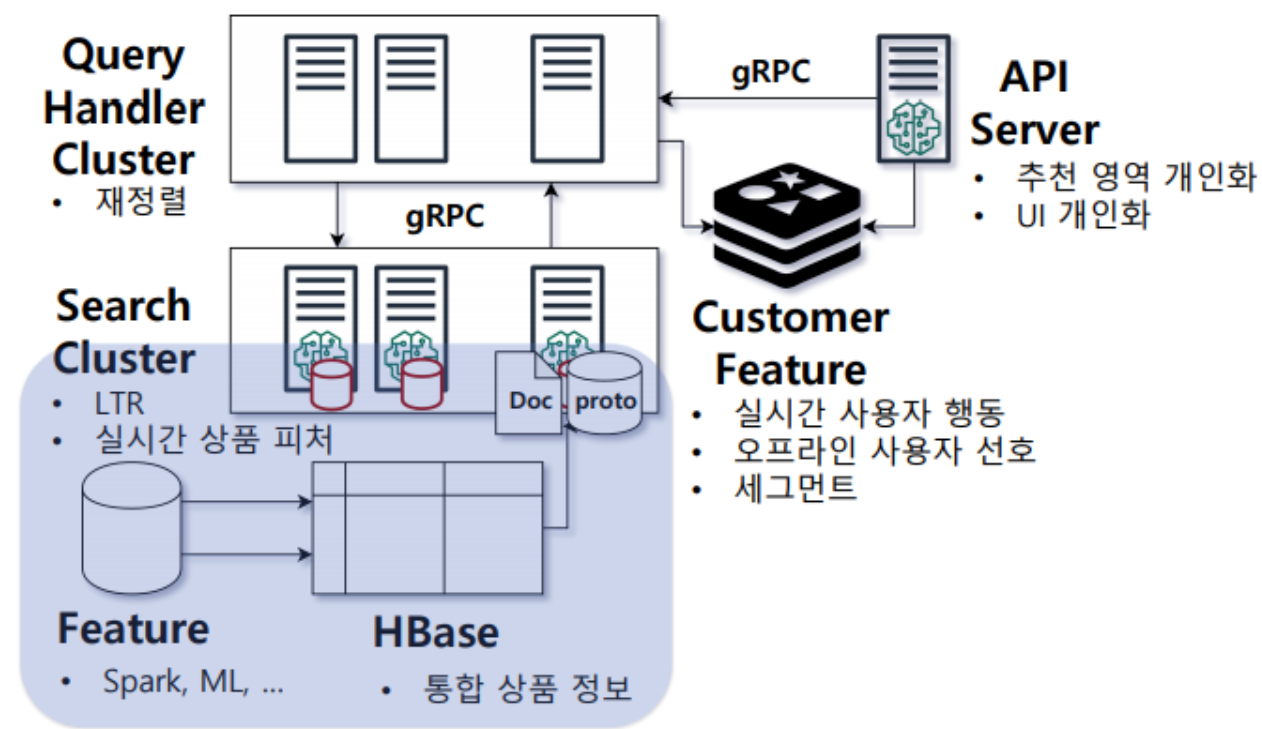
빨래하기좋은날 빨래
대형, 1개

14,560원

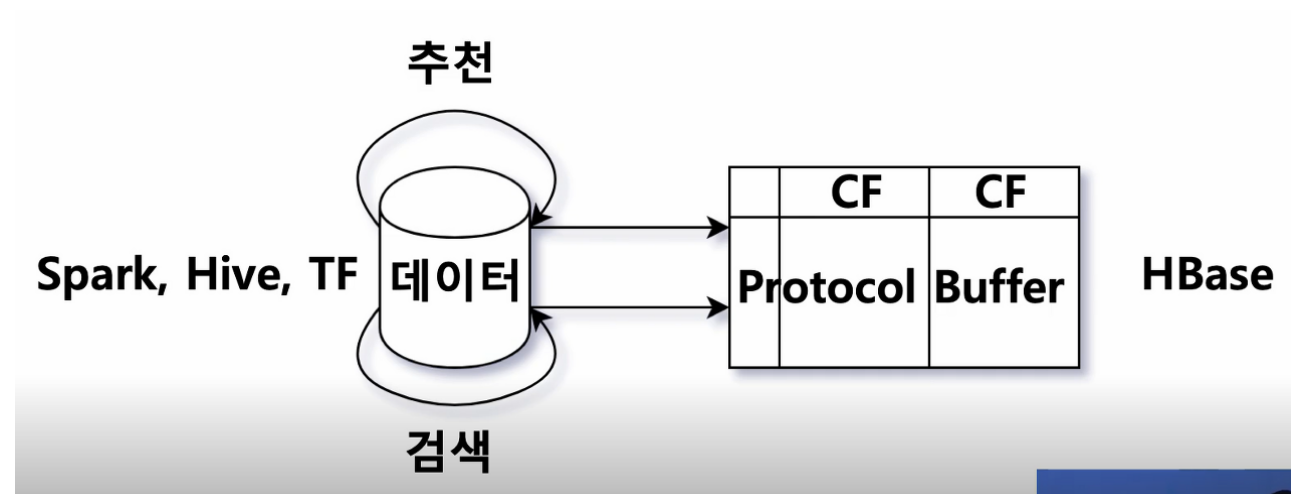
로켓배송

★★★★★ (8,071)

현재 추천 시스템 탐구



전체 시스템은 위와 같이 생겼다.



모델은 Spark, Hive, TF 등을 활용해 데이터의 Feature를 추출한다. 추출한 Feature는 Protocol Buffer(구글에서 만든 직렬화(Serialization)된 데이터 구조)의 형태로 HBase에 저장한다.

쿠팡의 Protocol Buffer

```

message Product {
    int64 product_id = 1;
    Image image = 2;
    Category category = 3;
    int32 rating = 4;
    SearchSignals search_signals = 5;
    RecoSignals recommendation_signals = 6;
}
message RecoSignals {
    repeated RecoSource viewed_together = 1;
    repeated RecoSource bought_together = 2;
}
message RecoSource {
    int64 product_id = 1;
    double score = 2;
}

```

* Product라는 메시지를 정의해서 관리

* 상품 그 자체가 갖는 feature인 **image, category, rating**, 리뷰 수 등의 정보 있음

* 이 상품이 다른 어떤 상품에 몇 점의 점수로 추천 할 수 있는지를 반복 필드(?)로 두고 있다.

* search팀에서 관리하는 정보, 추천팀에서 관리하는 정보가 서로 다른 column family에 들어있기 때문에 인덱싱 할 때에는 Protocol Buffer의 mergeFrom API를 사용해 merge해서 사용한다.

Search Cluster에서는 컨텍스트(유저, 아이템, 시간, 카테고리)와 관련된 상품을 찾고, 조건에 따라 필터를 한 다음 점수에 따라 정렬한다.

근데 모델에서 만드는 feature의 수는 계속해서 많아진다. 이럴수록 **어떤 feature에 얼마만큼의 가중치를 줘야하는지 결정해야 하는 문제**가 생긴다. 이러한 작업을 사람의 손에 맡기는 것은 효율성이 낮다.

이 작업을 컴퓨터가 할 수 있도록 쿠팡에서 도입한 것이 **LTR**이다. 즉, 쿼리 튜닝을 컴퓨터에게 맡기기 위해 사용하는 것이다.

LTR을 학습시키기 위해서는 어떤 상품이 어떤 feature를 가지고 추천되었는지 알아야 하며, 반응이 어땠는지도 알아야 한다. 그래서 쿠팡의 서버에서는 추천 상품을 내보낼 때 마다 해당 상품이 어떤 feature를 가지고 있었는지 logging한다. 또한 해당 상품이 어떤 session id에, 어떤 search id에 추천되었는지도 기록되어야 한다.

컨텍스트를 정의하는 것이 **Query Handler Cluster**다.(=서비스를 정의한다) 예를 들어, “냄비와 함께 살 만한 할인상품”이라는 추천을 하고 싶다면, bought_together에 “냄비”가 들어있는 상품을 Query하고, category를 “식품”으로 필터링한다. 또한 discount_rate로 부스팅 까지 하면 된다. 이 때 Customer Feature를 사용해서 개인화된 서비스를 정의하기도 한다.

Feature Selection

모든 피처를 사용하는 것이 성능이 더 좋을 수는 있다. 하지만 쿠팡에서는 피처의 수를 적게 유지하도록 노력한다고 한다. 그 이유는 피처가 많

다면 새로운 피처를 넣었을 때 어떻게 영향을 미치는지 파악하기 어렵고, 어떤 것을 먼저 개선해야 하고, 디버깅해야 하는지 파악하기가 힘들다.

좋은 feature를 선택하기 위해서 모델이 어떤 feature를 중요하게 여기는지 분석하기도 하고, feature 간의 상관관계를 따져 상관관계가 높은 feature는 삭제한다.

Raw Feature vs Processed Feature

쿠팡은 Processed Feature 사용

Raw Feature를 사용하면 각 feature가 모델에 어떻게 영향을 미치는지는 비교적 쉽게 이해할 수 있지만, 상관관계가 높은 feature가 여러 개 남을 수 있다고 한다. 처음 LTR 모델을 도입할 때는 Raw feature를 사용하되 시간이 지나면 상관관계가 높은 feature를 제거한 processed feature를 사용하는 것이 바람직하다.

복잡한 LTR 모델(RNN, CNN) vs 단순한 LTR 모델(Decision tree, MLP)

쿠팡은 굉장히 단순한 모델 채택(Decision tree)

이유

1. 각 픽셀의 RGB 값이 모두 feature가 되는 이미지와 달리 상품 추천 영역에서는 feature의 수가 굉장히 적은 편이다.
2. 복잡한 모델을 사용해서 오프라인 평가를 좋게 받더라도 유저의 반응이 그 만큼 좋으리라는 보장이 없다.

피처 엔지니어링 vs 파라미터 튜닝

쿠팡은 피처 엔지니어링 선호

파라미터 튜닝을 이것저것 시도를 해보고 있지만, 어떤 파라미터가 왜 잘 작동하는지 이해하기 어렵다. 그보다는 새로운 feature를 만들고, 그 feature가 모델에 어떤 영향을 미치는지 파악하는 것이 중요

Q&A

1. 새로 들어온 유저에 대해서는 어떻게 대응하는가?

들어와서 상품을 보기 시작하면서 정보가 쌓이기 시작한다. 이것이 최종적으로 유저의 선호를 결정하는 것은 아니기 때문에 초기 유저의 선호는 자주 바뀐다. 따라서 추천되는 상품 등도 자주 바뀌게 된다.

2. 추천 시스템 평가 방법

서비스 개발 주기가 굉장히 빨라서 offline 평가는 크게 신경쓰지 않는다. A/B테스트를 바로 진행하며, 그 결과로 의사결정을 한다.

3. DBMS는 어떤 것 사용하는가?

분산 HDFS(Kafka)에 저장 => 하이브에 저장 => 스파크나 하이브SQL 등 작업자가 선호하는 방법으로 꺼내 쓴다.

4. A/B테스트 검증은 어떻게?

p-value

5. A/B테스트 샘플군은 어떻게 정하나?

정하지 않는다. 전체 유저를 반으로 잘라 테스트한다.

6. A/B테스트를 하는 타이밍

모델이나 feature를 하나 만들고 난 후 즉시 수행. 성능이 안 좋으면 바로 폐기

7. 유저가 들어올 때 마다 매번 같은 추천 상품을 노출하는가?

No. random 함수를 사용해 매번 다른 상품 노출

1

구독하기



이름

암호

☐ Secret

여러분의 소중한 댓글을 입력해주세요.

댓글달기

하루배송, 새벽배송 등 배송 속도에 열을 올렸던 이커머스 업체들이 맞춤형 할인쿠폰과 상품 추천 등 개인화 서비스에 집중하고 있다. 맞춤형 할인쿠폰은 상대적으로 적은 마케팅 비용으로 구매 도달률을 높이는 타겟 마케팅 효과가 크고, 개인화 추천 서비스는 실질적인 매출 증대 효과가 있단 게 업체들의 설명이다.

29일 이커머스 업계에 따르면 지난해부터 쿠팡은 문자메시지 수신에 동의한 이용자를 대상으로 맞춤형 할인쿠폰을 발급하고 있다. 예를 들어, 쿠팡 웹 페이지나 앱에서 가전제품을 검색한 이력이 있거나 가전제품을 주문하지 않고 장바구니에 담아뒀다면 개인화 서비스 일환으로 이들에게만 가전제품 할인쿠폰을 문자메시지를 통해 지급한다.

할인쿠폰은 쿠팡이 제시한 웹주소로 들어가 로그인을 해야만 받을 수 있으며, 유효기간 만료 후엔 재발급이 불가능하다. 쿠팡 카카오톡 플러스친구나 쿠팡 이벤트쿠폰 페이지 등 타 마케팅 채널로는 지급되지 않는다. 최근엔 삼성전자 에어드레서나 건조기, 공기청정기 등을 구매 시 쿠팡 판매가에 5%를 추가로 할인 받을 수 있는 `삼성전자 5% 할인쿠폰`이 일부 이용자에게 발급됐다. 이전엔 쿠팡캐시 이벤트 등이 열렸다.

쿠팡에 따르면 문자메시지 마케팅은 이용자 빅데이터를 활용해 특정 발송 기준(로직)에 따라 실시하는 단발성 이벤트로, 특정 기간에 한해 운영한다.

이 같은 문자메시지 타겟 마케팅에 대표적인 곳은 마켓컬리다. 마켓컬리는 장바구니 쿠폰, 특정상품 할인쿠폰 등을 개인 또는 단체에게 시시때때로 지급한다. 앞서 일부 이커머스 업체들도 장기미사용고객(휴먼고객)이나 신규고객 등을 대상으로 할인쿠폰을 지급해왔지만, 개인별 할인쿠폰을 주된 마케팅 방식으로 활용하는 것은 마켓컬리가 선도적이었다. 이에 따라 문자메시지를 통한 마켓컬리 앱 유입률도 꾸준히 상승하고 있다.

마켓컬리 관계자는 "많은 쿠폰을 발행한다고 해도 해당 이용자에게 필요한 쿠폰이 아니라면 결국 서비스를 이용하지 않는다"며 "유사한 성향을 가진 고객 집단을 분류해 맞춤형 할인쿠폰 등 개인화된 혜택을 제공하고 있다"고 설명했다.

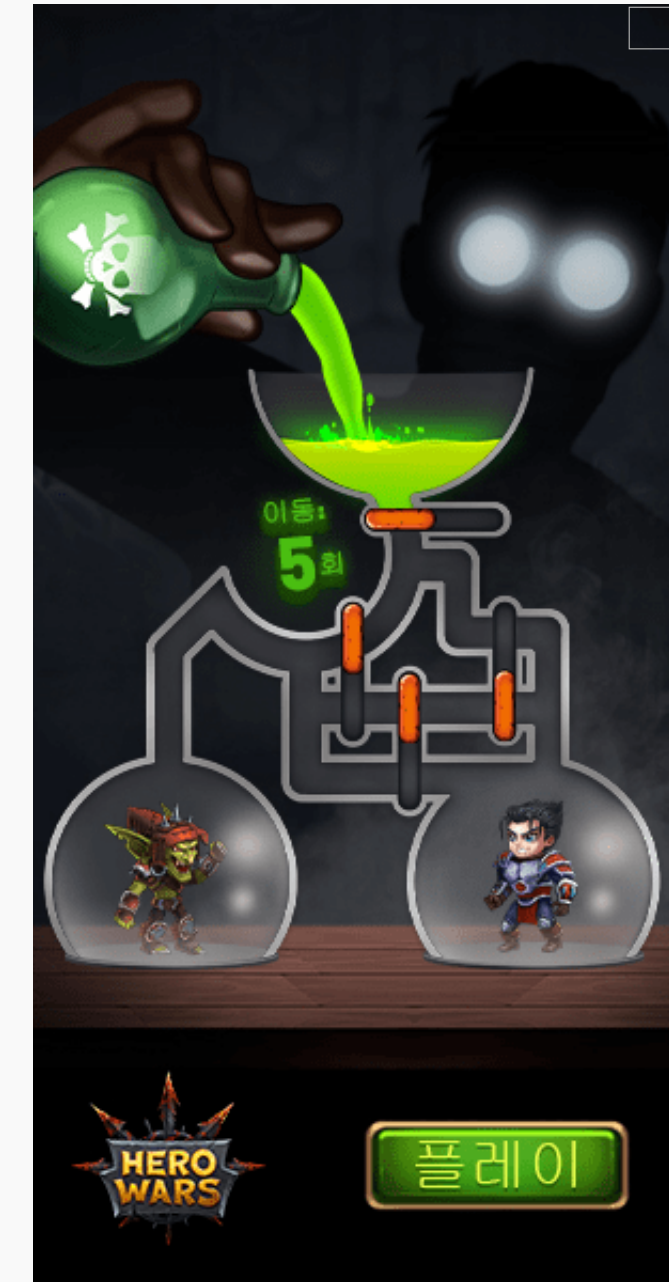
티몬도 최근 딥러닝 기술을 적용한 개인화 추천 서비스 `포유(For You)`를 내놔다. 구매 습관과 패턴, 클릭 이력에 따라 이용자를 17가지 취향으로 나눠 14가지 로직을 통해 상품을 추천한다. 상품별 구매주기를 분석해 재구매 알림 서비스도 제공한다. 서브 메뉴가 아닌 단독 메뉴로 구성해 비중을 높였으며, 올해 하반기부터 본격 적용한다.

앞서 네이버 역시 인공지능(AI)기반의 쇼핑추천기술인 에이아이템즈(AiTEMS)를 활용해 개인의 관심사와 선호도에 따라 상품을 추천하는 포유(FOR YOU) 시범 서비스를 내놔다. 네이버에 따르면 지난해 에이아이템즈 고도화 이후 일평균 클릭수가 6개월만에 약 27% 증가했으며, 올 상반기 내 확대적용할 예정이다.

11번가는 이용자가 구매할 것으로 예상되는 상품을 알고리즘을 이용해 가장 살 것 같은 시간에 모바일 알림을 보내는 리마인드




허쉬초콜릿 호빵도 배달...GS25, 호빵 출시

`고깃집도 맛으로 승부하니 백화점 온라인몰..



인기뉴스

종합 연예 스포츠

- 1 7만원 나오던 관리비가 1297만원으로?... 대전 아파트 논.. 
- 2 北, 보란듯이 총살하고 시신 불태웠는데... 손쓰지 못한.. 
- 3 신동빈의 집념...시애틀에 美 세번째 호텔 

서비스를 운영한다. 유사한 제품을 이미 구매했거나 최근에 장바구니 리마인더 알림을 받았다면 알림 메시지를 보내지 않아 앱 알림 피로도를 낮춘 것이 특징이다. 장바구니 상품이 여러 개일 경우 이용자가 가장 매력을 느낄 상품만 로직을 통해 먼저 추천한다. 11번가에 따르면 올해 들어 지난달까지 `고객 맞춤형 장바구니 알림` 서비스를 시범 운영하면서 월 평균 10억원이 추가로 결제되는 등 개인화 서비스가 효과를 보고 있다.



이커머스 업계는 이용자의 친밀도와 쇼핑 편의성을 높인단 점에서 개인화 서비스가 더욱 확대될 것으로 보고 있다. AI와 딥러닝, 스마트러닝 등 최신 IT기술 적용 사례도 늘고 있다. 다만, 일각에서는 자칫 정보와 서비스 편향 가능성이 있단 우려가 나온다. 개인화 서비스가 확대되는 만큼 피싱 주의도 요구된다.

이커머스 업계 관계자는 "개인별 할인쿠폰의 경우 업체는 마케팅 비용을 줄일 수 있고 이용자는 나만의 특별한 서비스를 받는단 기분을 느낄 수 있지만 특정 소비자에게만 할인쿠폰이 지급돼 일부 반발도 가능한 만큼 타겟팅에 신중할 필요가 있다"고 말했다.

[디지털뉴스국 배윤경 기자]

[© 매일경제 & mk.co.kr, 무단전재 및 재배포 금지]

#쿠팡 #티몬 #마켓컬리 #11번가 #네이버

오늘의 핫 뉴스



보란듯 총살하고 시신 불태웠는데...손 못쓰는 軍 12시간 32분

리니지보다 더 번 `배그`..."돌격! 코스피로" 12시간 28분

"진짜 없어요. 돈 줘도 못들어가요"...하남 전세대란 19시간 11분

금융감독원장의 엄포에도...은행 점포 내달 40곳 폐쇄 12시간 49분

주식으로 "돈 다 까먹는 이유", 언젠간 오르겠지... AD

코로나 덮친 시장 가보니 "대목인데 상인이 더 많아" 12시간 52분

4 당선 2주만에...더시민당, 양정숙 제명 추진 왜?



5 [매부리TV] "진짜 없어요 돈 줘도 못들어가요"...3기 신..



AD 소노호텔앤리조트(대명) 코로나19 경기침체 한시적 혜택!...



가장 많이 공유된 기사

f 506 "명동밥집에서 희망 한 그릇 드시..

f 399 AI로 피자 구워요...직원 20%가..

f 364 中 바이러스 연구소 `양산 유치설..

f 258 [종합] `패트 총돌` 첫 공판...

f 239 이번엔 배터리 혁신...테슬라의 `소..

f 218 자녀 이야기에 `발끈`한 추미애...진..

f 216 [Mobile World] 디지털..

f 189 英 조앤롤링 얘기꺼낸 이낙연 "해..

f 174 집주인 "입주 못하고 세금폭탄"...

f 163 [핫이슈] "나랏돈이니 돈이나"..

포토