**비즈니스 애널리틱스 사례발표 자료조사**

**[AiTEMS(에이아이템즈)]**

**옥영** #기존 URL 내용 정리

1. AI·빅데이터로 맞춤상품 추천… 입소문 타고 90% 이상 예약

(<http://www.dt.co.kr/contents.html?article_no=2020083102100551054001&ref=naver>)

또한 네이버는 이용자의 개인행동 데이터를 기반으로 개별 소비자의 관심사나 취향을 분석해 상품을 추천하는 AI 시스템인 'AiTEMS(에이아이템즈)' 기술을 고도화해 판매자에 제공하고 있다. 네이버측은 "2017년 9월 오픈 이후 에이아이템즈 일일 클릭량이 4개월만에 7배 이상 증가하는 등 개인화 상품 추천 서비스에 대한 이용자 만족도가 지속적으로 올라가고 있다"면서 "동시에 이용자 취향에 따라 인지도 낮은 상품 등도 골고루 노출되는 기회를 얻으면서 판매자 만족도도 상승했다"고 설명했다.

1. 네이버쇼핑, 빅데이터 활용한 선호 상품 선별 시스템 도입

(<http://newsbiz.co.kr/news/newsview.php?ncode=179524781323873>)

네이버가 AI기반의 상품 추천 시스템 AiTEMS(에이아이템즈)를 활용해 검색 결과에서도 개인의 관심사, 선호도에 따라 쇼핑을 경험할 수 있는 ‘FOR YOU (베타)’ 영역을 신설했다고 밝혔다.

‘FOR YOU’는 네이버 검색 결과에서 전체 이용자 대상으로 동일한 결과를 제공하는 ‘전체상품’ 영역 옆에 추가로 신설되는 영역으로, 로그인 기반의 사용자가 선호할만한 몰(mall)과 상품을 추천함으로써 개인에 보다 특화된 쇼핑 결과를 제공한다.

FOR YOU 영역은 ‘패션 의류 및 잡화’ 카테고리의 일부 품목(원피스, 코트, 롱패딩, 가디건 등)에 선적용되며, 최근 쇼핑 이력이 많은 사용자 대상으로 노출된다. 예를 들어, 네이버 모바일 검색창에 ‘원피스’를 입력하면 FOR YOU영역을 통해 자주 방문 및 클릭했거나 이전에 구매했던 몰, 선호할만한 몰의 리스트와 해당 몰의 인기 원피스 상품을 한 눈에 확인할 수 있는 식이다.

또한, FOR YOU 영역 더보기를 클릭하면 선호할만한 몰과 유사한 몰들을 한데 모아 보여주는 ‘연관몰 추천’ 영역을 통해 내 취향에 맞는 다양한 쇼핑 아이템들을 더욱 다채롭게 만나볼 수 있다.

이를 위해 네이버는 현재 네이버 모바일 쇼핑 서비스에 적용하고 있는 AiTEMS 기술을 쇼핑 검색에도 활용하여, 사용자 정보(▲성 ▲연령 등)와 쇼핑 이력(▲구매 ▲찜 ▲클릭 ▲검색 ▲장바구니 등)을 분석해 맞춤형 검색결과를 제공하는 연구를 진행했다.

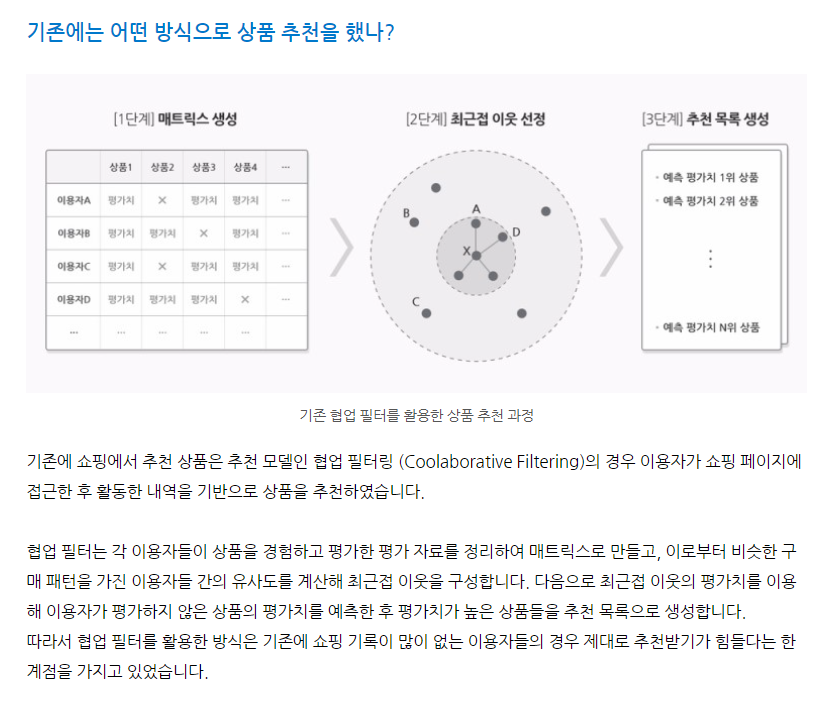
특히 ‘패션’ 카테고리에서 나타나는 소비 패턴을 반영해, 쇼핑몰, 브랜드, 스타일 태그, 카테고리 위주의 이력을 중점적으로 분석해 결과에 반영했다.

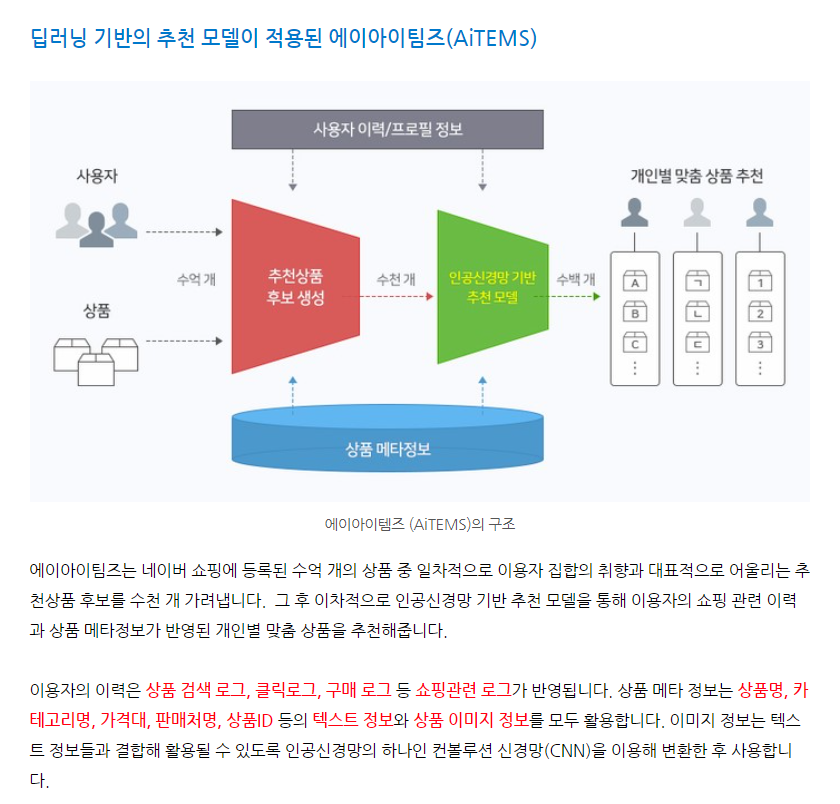
향후 네이버는 FOR YOU를 가전, 리빙, 식품 등 다양한 카테고리와 브랜드로 확대 적용하며, 개인화 추천 영역을 강화해나갈 예정이다.

네이버 AiTEMS 이정태 리더는 “지난해 6월 AiTEMS 고도화 이후, AiTEMS 추천 영역에서 발생하는 일평균 클릭수가 6개월만에 약 27% 증가하는 등 사용자와 판매자의 만족도가 지속적으로 올라가고 있다”며, “AiTEMS를 연구하며 쌓아온 기술과 서비스적 노하우를 검색에도 활용함으로써, 사용자에게는 더욱 편리한 쇼핑 경험을, 판매자들에게는 더 많은 노출로 성장 기회를 제공할 수 있기를 바란다”고 밝혔다.

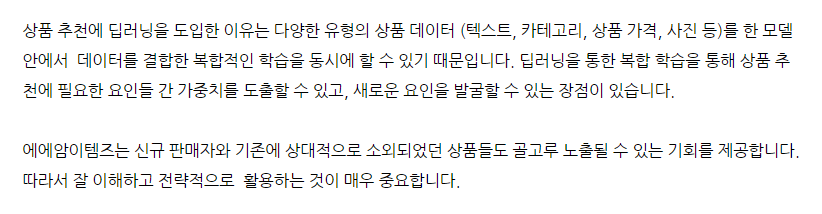
1. 딥러닝 에이아이템즈(AiTEMS) 추천상품 노출 알고리즘, 모바일 메신저 쇼핑몰 선점 효과?

(<https://blog.naver.com/PostView.nhn?blogId=hksns&logNo=221533073475&parentCategoryNo=&categoryNo=6&viewDate=&isShowPopularPosts=true&from=search>)









**정기적으로 구입하는 상품 재추천, 여러 카테고리의 추천 상품을 한눈에, 비슷한 취향의 브랜드와 쇼핑몰 추천 등이 있음.**

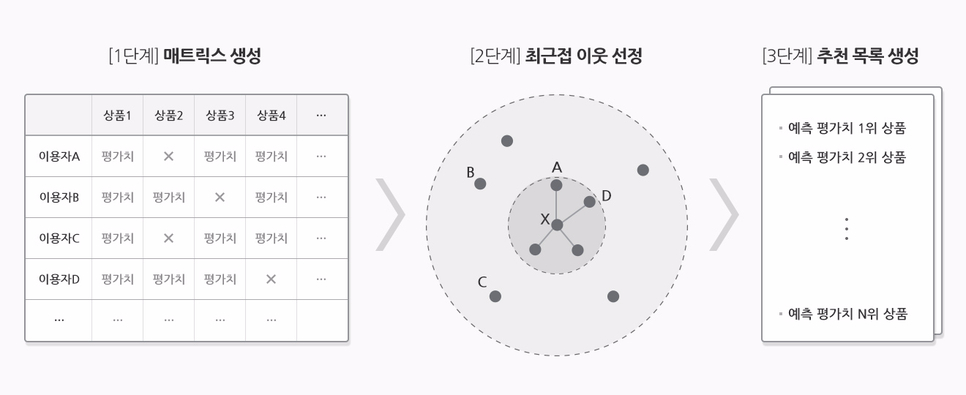
1. 취향저격 잇템 찾아주는 딥러닝 기반 개인화 상품 추천 시스템 ‘에이아이템즈 (AiTEMS)’ (<https://blog.naver.com/naver_search/221085553045>) – 네이버 공식 블로그

* 에이아이템즈가 도입되면 쇼핑 추천이 어떻게 바뀌게 되나요?

: 에이아이템즈의 추천 결과는 **모바일 첫화면 쇼핑판의 ‘AiTEMS 추천’ 섹션**에 보여지며, **쇼핑MY 페이지**에서도 확인할 수 있습니다.

* 기존에는 어떤 방식으로 쇼핑 추천이 이루어졌었나요?

: 국내외 많은 쇼핑 사이트들에서 활용되고 있는 추천 모델인 **협업필터링 (Collaborative Filtering)**의 경우, 이용자가 쇼핑 페이지에 접근한 후 활동한 내역들을 기반으로 상품을 추천해 주고 있습니다.



협업필터는 각 이용자들이 상품을 경험하고 평가한 평가치들을 정리하여 매트릭스를 만들고, 이로부터 비슷한 구매 패턴을 가진 이용자들 간의 유사도를 계산해 최근접이웃을 구성합니다. 다음으로 최근접이웃의 평가치를 이용해 이용자가 평가하지 않은 상품의 평가치를 예측한 후, 평가치가 높은 상품들을 추천 목록으로 생성합니다.

협업필터를 활용한 방식은 기존에 쇼핑 기록이 많이 없는 이용자들의 경우 제대로 추천을 받기 힘들다는 한계점을 가지고 있습니다.에이아이템즈가 어떻게 이 문제를 개선했는지에 대해서는 다음 포스팅에서 다루도록 하겠습니다.

1. (다음 포스팅) 쇼핑을 자주 하지 않는 이용자도, 신규 판매자도 만족스러운 쇼핑 플랫폼을 만드는 에이아이템즈(AiTEMS)의 기술 (<http://naver.me/5hC2PkoS>) – 네이버 공식 블로그

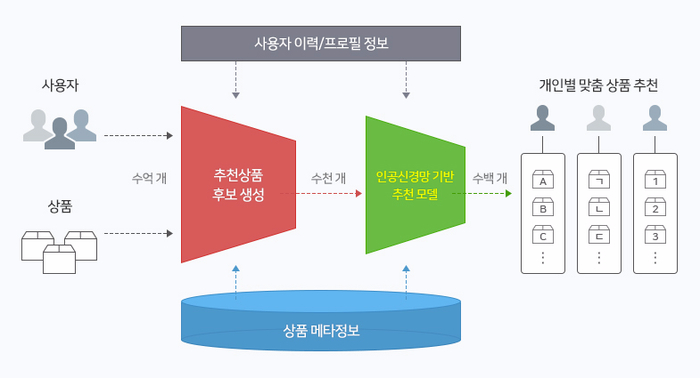
* 에이아이템즈만의 차별점은 무엇인가요?

현재 네이버에 등록된 쇼핑 상품의 개수는 5억여 개에 달하고 매일 400만 여개의 신규 상품이 등록되고 있지만, **검색 채널을 통해서는 이용자들이 많이 찾는 인기상품이 상대적으로 더 잘 노출되는 경향이 있습니다.**

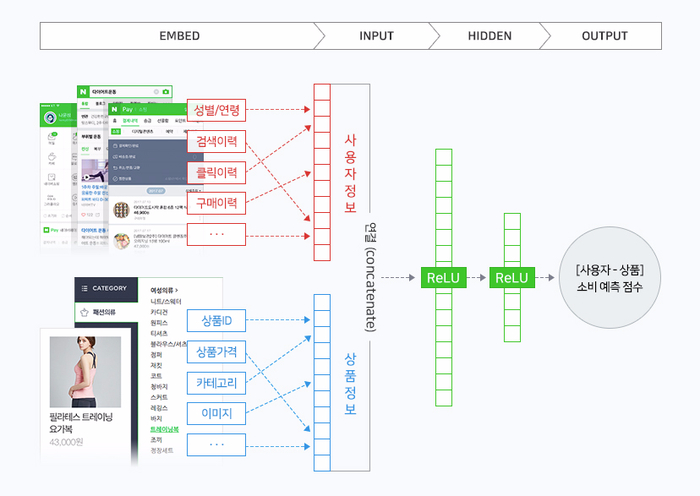
에이아이템즈는 이용자의 취향을 파악하고, 상품 자체가 가지고 있는 취향저격 요소와 매칭함으로써 보다 더 다양한 상품들이 추천될 수 있도록 개발되었습니다. **인지도가 없는 신규 상품일지라도 상품이 가지고 있는 특성을 분석해 이용자 취향에 맞는 요소가 포함되어 있을 경우, 인기상품과 동일하게 추천 대상이 됩니다.** 그렇기 때문에 신규로 등록한 판매자의 상품들도 골고루 유통되는 효과를 기대할 수 있습니다.

* 에이아이템즈에는 어떤 기술이 활용되었나요?

인공지능을 뜻하는 AI와 상품을 뜻하는 Items의 합성어인 에이아이템즈에는 딥러닝 기반의 추천 모델이 적용되었습니다.



에이아이템즈는 우선 수억 개의 상품 중 일차적으로 이용자 집합의 취향과 대표적으로 어울리는 추천상품 후보들을 수천 개 추려냅니다. 그 후, 인공신경망 기반 추천 모델을 통해 이용자의 쇼핑 관련 이력과 상품 메타정보가 반영된 개인별 맞춤 상품을 추천해주는 방식으로 작동합니다. **이용자의 이력**에는 **상품 검색/클릭 로그, 구매 로그** 등 쇼핑 관련 로그가 반영됩니다. **상품 메타 정보**의 경우 **상품명, 카테고리명, 가격대, 판매처명, 상품ID 등의 텍스트 정보와 상품 이미지 정보를 모두 활용**합니다. 이미지 정보는 텍스트 정보들과 결합해 활용될 수 있도록 인공신경망의 하나인 컨볼루션 신경망(CNN)을 이용해 변환한 후 사용합니다.



에이아이템즈의 핵심인 인공신경망 기반 추천 모델에서 이용자의 쇼핑 관련 정보, 상품 정보와 같은 서로 다른 형태의 입력 정보들은 먼저 **가상 공간 위에 표현 가능한 고차원의 벡터로 변환 (‘임베딩’)** 되고 이를 모두 병합하여 입력 벡터를 생성합니다. 그 후 기존에 **학습된 은닉층 (hidden layer) 내 가중치 행렬 연산과 활성화 함수 (ReLU; Rectified Linear Unit)**를 통과하면서, 이용자가 상품을 구매할 확률이 얼마나 될지 예측하는 **소비 예측 점수**가 도출되게 됩니다.

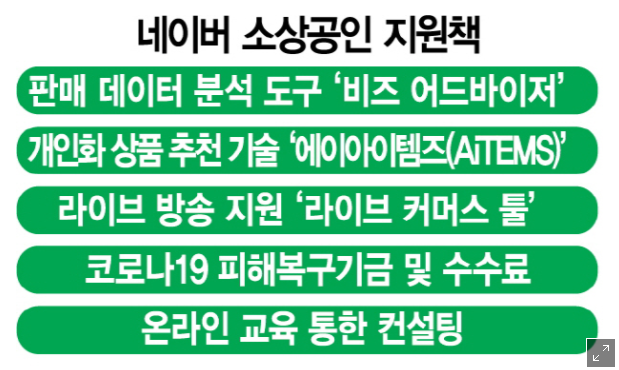
* 쇼핑 추천에 딥러닝을 도입한 이유는 무엇인가요?

상품 데이터는 다양한 형태의 정보를 갖고 있습니다. **딥러닝 모델을 활용하면 텍스트, 숫자, 이미지와 같이 유형이 다른 데이터들을 한 모델 안에서 동시에 학습할 수 있기 때문**에, 상품 카테고리, 가격, 사진과 같은 서로 다른 형태의 데이터를 결합한 복합적인 학습이 가능해집니다. 이러한 복합적 학습을 통해 추천에 필요한 요인들 간 가중치를 기계가 스스로 알아낼 뿐만 아니라, 입력 요인들을 조합해 새로운 요인을 발굴해 내기도 합니다.

딥러닝을 활용한 상품 추천 시스템 에이아이템즈는 패션 분야 쇼핑에 먼저 도입된 후, 리빙 등 다른 분야로도 점차 확대 적용될 예정입니다. 쇼핑윈도의 경우 각 윈도에 특화된 추천 방식 적용도 검토하고 있는데요. 에이아이템즈는 **신규 판매자와 기존에 상대적으로 소외되던 상품들도 골고루 노출될 수 있는 기회를 제공하는 <프로젝트 꽃>의 의미를 기술로 실현해** 나갈 예정입니다.

**주미**

1. **포스트 코로나 : 네이버, 스마트 스토어로 700만 소상공인 ‘디지털 전환’ 돕는다**

코로나19 장기화로 전 세계 소비 트렌드가 급격하게 온라인으로 전환되는 가운데 네이버가 소상공인들의 ‘구원투수’로 등판했다. 네이버는 누구나 쇼핑몰을 개설할 수 있는 자체 플랫폼인 ‘스마트스토어’와 매출을 증대시킬 수 있는 기술을 함께 지원하면서 국내 700만 소상공인의 ‘디지털 트랜스포메이션(전환)’을 돕고 있다.  
  
‘언택트(비대면)’ 문화의 확산은 비즈니스 환경에도 중대한 변화를 가져왔다. 한 데이터 분석 결과에 따르면 코로나19가 확산한 3월 말을 기점으로 전 세계 온라인 쇼핑시간이 평균 47% 증가한 것으로 나타났다. 한국 역시 온라인 쇼핑시간이 29% 증가해 코로나19로 타격을 입은 오프라인 판매자들에게 사업 패러다임 변화가 절실한 것으로 나타났다. 디지털 전환이 오프라인 판매자의 생존을 가르는 문제가 된 셈이다.  
  
이러한 추세는 네이버쇼핑에서도 여실히 드러난다. 네이버는 지난해 7월부터 올해 6월까지 최근 1년간 연매출 1억원 이상을 달성한 판매자가 2만6,000명으로 전년 동기 대비 40% 이상 늘어났다고 밝혔다. 6월 한 달을 기준으로 봤을 때 매출이 1억원 이상인 판매자만 2,800명에 달한다. 스마트스토어를 적극적으로 활용해 디지털 전환에 성공한 사례다.  
이와 더불어 네이버는 스마스스토어 입점 소상공인에게 매출을 증대시킬 수 있는 기술 지원을 제공하고 있다. 판매 데이터 분석을 돕는 무료 도구 ‘비즈 어드바이저’와 개인화 상품 추천 기술 ‘에이아이템즈(AiTEMS)’가 대표적이다. 여기에 더해 최근 라이브 방송을 통해 물건을 판매하는 ‘라이브 커머스’ 시장이 급격하게 성장하면서 ‘라이브 커머스’ 툴도 지원해 새로운 기회를 제공하고 있다.  
  
네이버는 디지털 환경에 익숙치 않은 판매자들을 위한 온라인 교육에도 집중하고 있다. 네이버 파트너스퀘어에서는 ‘스마트스토어로 창업 준비하기’, ‘잘 팔리는 상품 상세페이지 만드는 법’, ‘스마트폰으로 쉽게 촬영하기’ 등 온라인 창업의 ‘A to Z’를 비대면 온라인 라이브로 교육하고 있다. 질의응답으로 이뤄진 ‘Q&A 찐라이브’, 판매자들의 사연을 중심으로 코칭해주는 ‘사심상담소’ 등 다양한 콘텐츠가 제공된다.  
출처 : https://www.sedaily.com/NewsVIew/1Z6NZSKKH5

1. **Aitems의 분석 의의**

네이버 쇼핑 ‘Aitesm’의 이용률이 최근 80% 넘어섰다

<https://m.etnews.com/20191017000175>

**“상품을 사용자의 맥락과 의도에 맞게 출시함으로써 출시 2년만에 ‘Aitems 이용률이 80%까지 확대되고, 거래액은 2배 이상 성장해 쇼핑 거래액의 10%를 넘어섰다”고 했다.**

<https://biz.chosun.com/site/data/html_dir/2019/10/31/2019103100874.html>

**먼저 상품추천 AI인 에이아이템즈(Aitems)는 출시 2년 만에 이용률이 80% 늘었으며** 에이아이템즈를 통한 **쇼핑 거래액의 비중은 10% 이상으로** 나타났다. 에이아이템즈는 매주 1천600만명이 방문하는 네이버쇼핑에서 주요하게 사용되는 기술이다. 에이아이템즈 30만명의 판매자가 등록한 8억개의 상품을 사용자의 성별과 연령대, 구매주기, 패턴 등에 기반해 쇼핑의 흐름을 이끈다.

한 대표는 “전체 쇼핑몰 중 80%가 에이아이템즈 판매자에게 실질적인 도움을 주고 있다”며 “모든 사용자를 만족시키고 보다 개인화된 쇼핑 경험을 제공할 수 있도록 AI를 확대 적용할 계획”이라고 설명했다.

쇼핑 성장에 힘입어 네이버페이의 **결제액은 4조원을 돌파했다**. 네이버는 9월 정식 출시한 테이블주문 기능을 시작으로 국내 네이버페이의 오프라인 진출을 확대할 계획이다.

한 대표는 “현재 주요 포스회사와의 협업을 통해 테이블주문이 빠르게 확장하고 있다”면서 “290만개 스몰 비즈니스 대상으로 특화된 솔루션을 적용해 나갈 것이고, 결제액 확대는 물론 양질의 리뷰 정보를 축적해 검색 서비스도 향상시킬 것이다”고 말했다.

아울러 한 대표는 쇼핑과 플레이스에서 일군 성공을 다음달 1일 네이버에서 분사하는 네이버파이낸셜에도 적용하겠다고 강조했다. 네이버파이낸셜은 네이버페이 사내조직(CIC)이 전신이다.

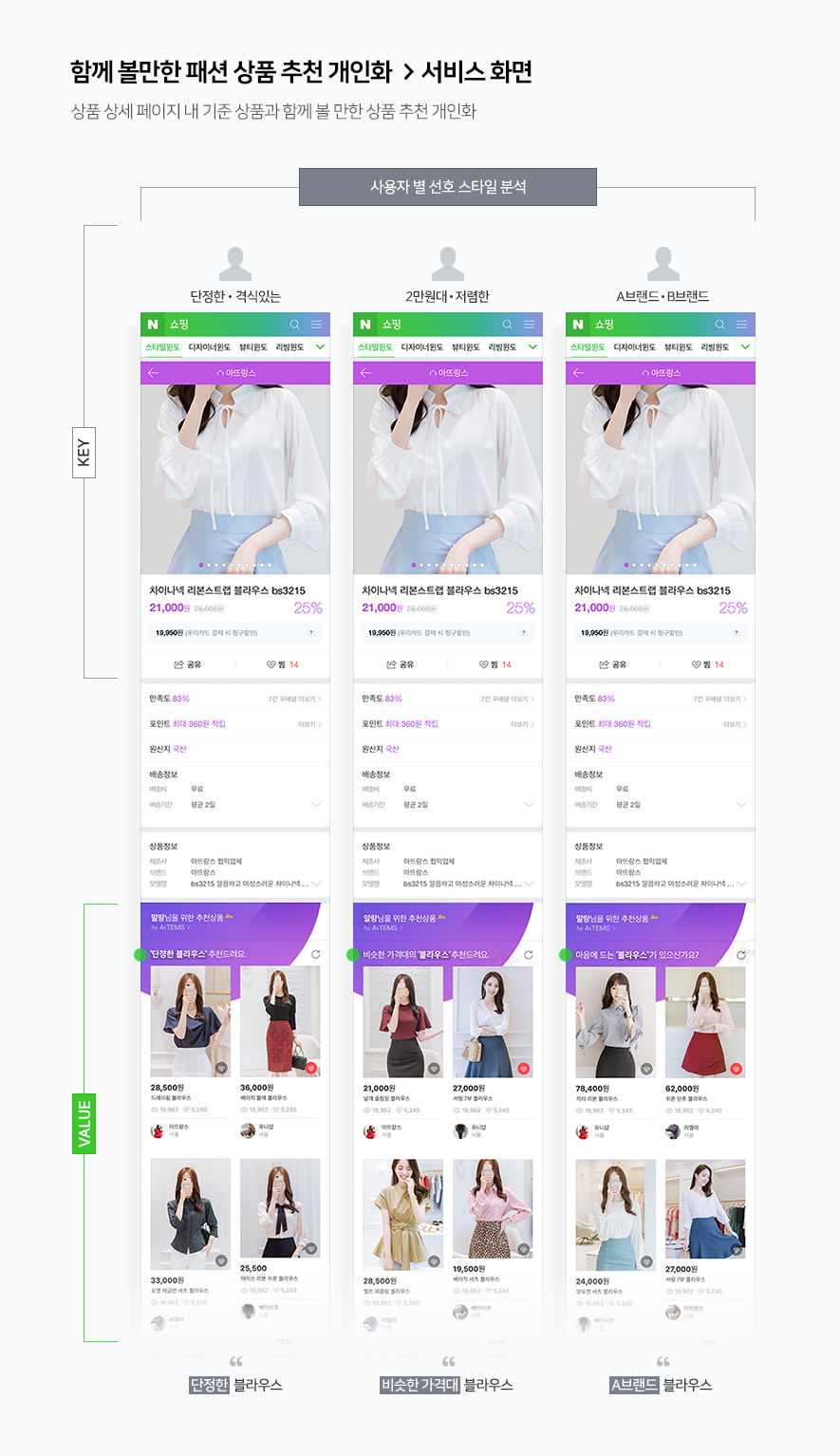
그는 “분사 후에도 결제규모 확대를 통한 금융 사업 기반을 키워 미래에셋과 새로운 시장에서 경쟁력 있는 상품을 만들겠다”고 밝혔다.

<https://zdnet.co.kr/view/?no=20191031101023>

1. **Aitems 기술 소개**

**어떤 기술이 활용되었을까?**

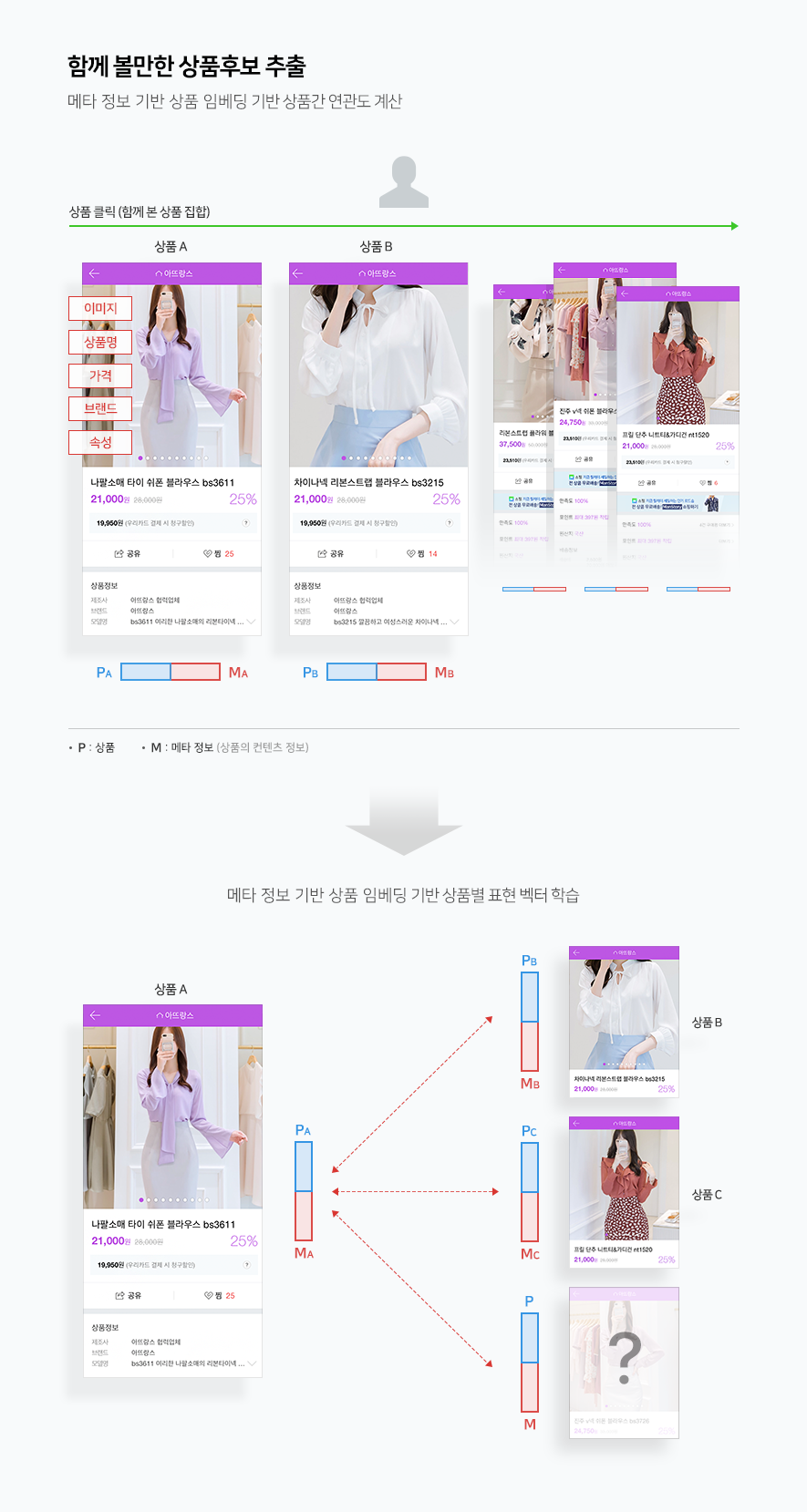
네이버 쇼핑에 스타일 토픽 기반 쇼핑 추천이 도입되었습니다. 백화점/아울렛/스타일 쇼핑 윈도 상품 페이지내에 **‘함께 볼 만한’ 상품 추천 개인화** 서비스가 새로 생기게 되었는데요.  
**특정 상품과 ‘함께 볼 만한’ 상품을 추출**하고, 이를 추천 후보로 두고 **사용자 스타일 선호도에 적합하도록 사용자마다 다르게**보여주게 됩니다.



**‘함께 볼 만한’ 상품 후보 추출**

지금까지의 ‘함께 볼 만한’ 상품 추천 서비스에서는 사용자들이 특정 상품과 함께 클릭해본(co-click) 상품을 추천해 주었습니다. 그러나 함께 클릭해본 상품이 충분하지 않은 경우가 대부분이므로 대다수의 상품에 대해서는 추천이 제대로 이루어지지 않는다는 한계가 있었습니다.  
그래서 상품의 컨텐츠 정보 간 연관성까지 활용하는 아래 예시 이미지와 같이 메타 정보 기반 상품 임베딩 모델을 기반으로 ‘함께 볼 만한’ 상품을 추출합니다.

**메타 정보 기반 상품 임베딩 모델**은 동일한 쇼핑 컨텍스트 내의 상품들 (한 사용자가 일정 시간 내 본 상품들)을 서로 ‘함께 볼 만한’ 상품 집합으로 가정합니다.  
하나의 쇼핑 컨텍스트 내 상품-쌍을 기반으로, word2vec 방식처럼 동일한 컨텍스트 내에서 공유된 **상품 고유 벡터 (P)**또는 **상품 메타정보 벡터 (M)**들은 서로 비슷한 값을 가지도록 학습됩니다. 여기서 메타 정보 (예: 상품 명 내 단어, 상품 브랜드, 상품 가격 등)는 상품 간에 공유되는 컨텐츠 정보로, 상품 간 일반적인 경향성을 학습하는 데에 활용되는 정보입니다.  
이후 개별 상품의 상품 고유 벡터와 상품 메타정보 벡터를 활용해 **상품 표현 벡터**가 구성되게 되며, 다른 상품에 대해서도 이와 유사한 방식으로 상품 표현 벡터가 구성됩니다. 그 다음 이러한 상품 표현 벡터들 간의 유사도 연산을 통해, 유사도가 높을 경우 ‘함께 볼 만한’ 상품으로 추천되게 됩니다. 여기서 메타 정보는 모든 상품에 부착되어 있기 때문에, 이를 통해 연관도를 계산하는 것이 가능하게 됩니다.  
이러한 시스템에서의 연산을 바탕으로, 각각의 상품 별로 일정 수준 이상의 연관도를 가진 상품들을 해당 상품의 **‘함께 볼 만한’ 상품 후보**로 간주합니다.



**스타일 토픽 기반 ‘함께 볼 만한’ 상품 추천 개인화**

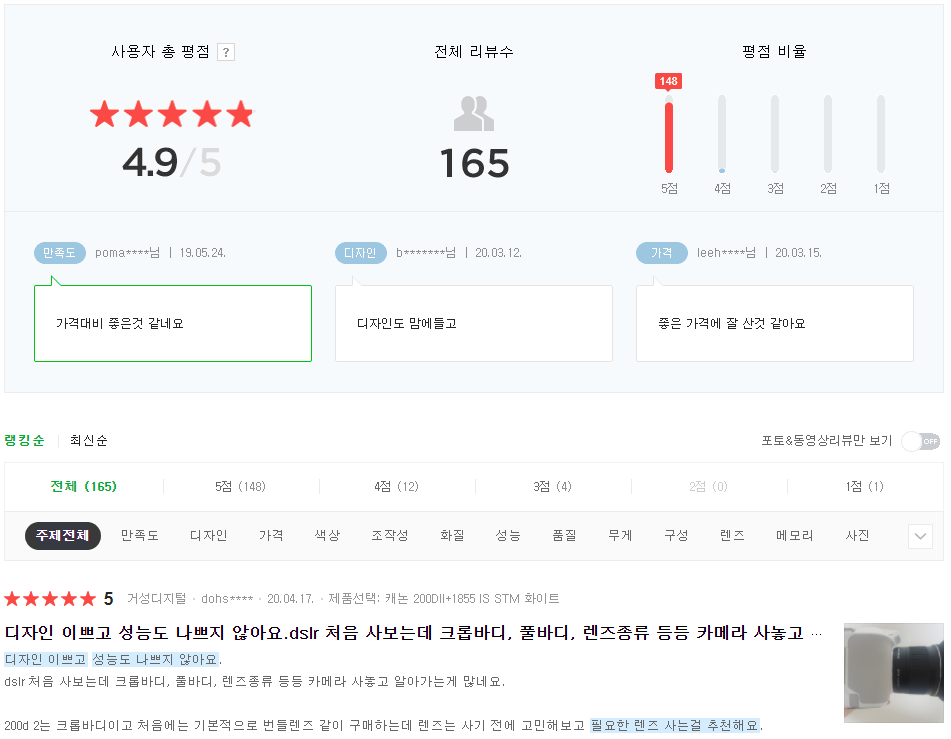
앞서 언급된 것과 같이 각각의 상품별로 추출된 ‘함께 볼 만한’ 상품들은 아래 예시 이미지와 같이 스타일 토픽 모델을 활용해 사용자 선호 스타일 토픽에 적합한 순서대로, 개인별로 다른 순서로 (랭킹) 노출되게 되는데요.



**스타일 토픽 모델**은 사용자가 선호하는 상품 집합 (클릭/구매/찜 등)이 그 사용자의 선호 스타일을 나타낸다는 가정에서 출발합니다.  
스타일 토픽 추출을 위해 상품의 상세설명/상품명/카테고리 정보 등의 텍스트 정보를 이용해 각 상품을 하나의 문서로 가정하고 토픽 모델링을 수행하게 됩니다.  
개인화 랭킹은 사용자가 선호하는 상품 집합을 이용해 **사용자의 스타일 토픽 분포**를 구하고, 앞서 언급되었던 **추출된 ‘함께 볼 만한’ 상품 후보들의 스타일 토픽 분포**를 구하여, 이들 간 유사도가 높으면 상품이 더 상위에 랭킹되게 되는 방식으로 작용합니다.

1. **네이버 쇼핑 리뷰에 적용된 감정 분석 AI**

네이버는 텍스트로부터 감정을 추출할 수 있는 감정 분석 기술을 2014년부터 쇼핑과 플레이스 영역에 적용해, 누구나 편리하게 리뷰 내용을 확인하고, 분석하고, 활용할 수 있도록 지원해왔습니다.

감정 분석이 적용되어 있는 서비스에서 이를 활용하고 계신 분들도 많을 텐데요, 아래 쇼핑의 사례처럼, **해당 제품에서 사용자들이 중요하게 생각하는 키워드(만족도, 디자인 등)들을 자동으로 추출 제공함으로써, 원하는 항목을 찾아볼 수도 있도록 지원하고** 있습니다. **또한, 리뷰의 내용 중에서 사용자들이 실제 사용 후 느낀 중요 부분에 파란색 음영(하이라이트)을 표기해 찾고자 하는 리뷰의 중요 내용을 보다 손쉽게 확인할 수 있습니다.**

**네이버 신규 감정 분석 모델만의 특징은 무엇인가요?**

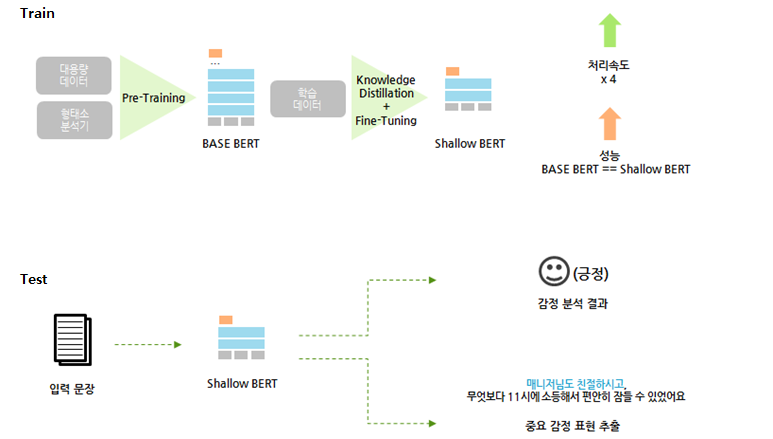
네이버의 신규 감정 분석 모델은 ▲ Base BERT 모델과 품질은 동등하면서도 4배 이상 빠른 처리 속도 ▲ 감정 분석(긍정/부정/중립) 정보뿐만 아니라, 감정 판단의 핵심이 되는 중요 감정 표현을 추출 및 제공 ▲ 쇼핑 리뷰, 호텔 리뷰, 맛집 리뷰, 뉴스 댓글, SNS 등과 같은 다양한 도메인에서 범용적으로 활용 가능하다는 점을 특징으로 들 수 있습니다.

​

그럼, 어떤 기술이 활용되었는지 살펴볼까요?

​

신규 감정 분석 모델에는 BERT(Bidirectional Encoder Representations from Transformer)라 불리는 모델이 사용되었습니다. CNN, LSTM 모델과 달리 빅데이터에서 얻을 수 있는 다양한 정보를 미리 학습(pre-training) 함으로써, 학습 데이터에 존재하지 않는 다양한 어휘적 특성을 활용할 수 있어 의미적 요소가 중요한 감정 분석 모델에 가장 적합한 모델입니다.



감정 분석에 최적화된 모델을 개발하기 위해 아래의 방법들을 적용하였습니다.

​

- 네이버가 보유한 빅데이터와 형태소 분석기를 활용해 한국어에 최적화된 방식으로 pre-training 을 수행하여 성능을 한층 향상시켰습니다. 또한, 서비스에 적합한 처리 속도를 얻기 위해 모델의 사이즈를 줄인 shallow BERT를 적용하여 처리 속도를 4배 이상 개선했으며, Knowledge Distillation 기술을 적용함으로써 Shallow BERT의 성능을 Base BERT 모델과 동등하도록 개선했습니다.

​

Knowledge Distillation은 주로 큰 모델의 지식을 효율적으로 추출해 작은 모델에 전달하여 품질을 향상하는 방법 중 하나입니다. 주로 딥러닝의 처리 속도 및 리소스 문제에 직면할 때 모델의 사이즈를 작게 줄이면서 품질 향상을 도모하기 위해 사용하는 방법입니다.

​

- 중요 감정 표현을 자동으로 추출하도록 모델을 설계했습니다. 일반적으로 중요 감정 표현을 추출하기 위해선 별도의 학습 데이터가 필요하지만 BERT의 Attention 정보를 이용하여 학습 데이터 없이 자동으로 중요 감정 표현을 추출할 수 있도록 설계/구축했습니다.

​

- 데이터 확장(data expansion) 및 전이 학습(transfer learning) 기술을 활용하여 효율적으로 신규 도메인을 확장할 수 있도록 설계해, 범용성을 확보​했습니다.

​

사용자들이 사용하는 단어와 표현들은 각 도메인마다 매우 차이가 큽니다. 예를 들면, “쇼핑 리뷰”에서는 “배송, 가격, 품질” 등의 상품 및 배송 관련 표현이 자주 나오지만, “뉴스 댓글”에서는 “집값, 정치, 임대” 등의 정치와 관련된 표현이 자주 나옵니다. 이러한 차이로 인해 "뉴스 댓글" 도메인만을 지원하는 감정 분석 모델을 그대로 "쇼핑 리뷰" 서비스에 적용하게 되면 만족스러운 품질이 나오지 않을 가능성이 매우 커집니다. 따라서 감정 분석 모델을 서비스에 적용할 경우 가장 중요한 요소 중의 하나는 최대한 다양한 도메인을 지원함으로써 충분한 범용성을 확보하는 것입니다.

​

신규 도메인을 확장할 경우, 대량의 데이터를 구축하는 것은 많은 시간과 비용이 소모됩니다. 이를 해결하기 위해 소규모 데이터를 이용하여 효율적으로 신규 도메인을 확장하는 방법을 설계/구축했습니다.

​

학습 데이터가 소규모인 신규 도메인 데이터를 확장하기 위해, 기존의 다른 도메인에서 학습된 모델을 가져와서 이를 기반으로 재학습하는 방법인 전이 학습의 특성을 활용하는 방법으로 공통점은 유지하면서 신규 도메인만의 특색 있는 데이터를 자동으로 추출 및 확장하도록 설계했습니다.

데이터를 확장하는 방법으로는 다수의 모델을 만들어서 다수결 원칙을 적용해 레이블(정답)이 없는 데이터에 대해 레이블(정답)을 자동으로 생성하여 데이터를 확장하는 “Ensemble” 방법을 기반으로 시스템 구축했습니다.

​

위의 방법으로 확장된 신규 도메인 데이터와 기존 데이터를 더해 신규 도메인을 추가로 지원하는 감정 분석 모델을 구축했으며, 전이 학습의 특성을 활용한 데이터 확장의 효과로 기존의 도메인들의 품질이 유지되거나 오히려 향상되는 효과를 얻을 수 있었습니다. 전이 학습의 특성을 활용하지 않는다면 도메인이 확장될 때마다 기존 도메인의 품질이 저하되는 현상이 발생합니다.

**신규 감정 분석 모델, 어떤 서비스에 활용할 수 있을까요?**

**쇼핑 영역**에서는 지금보다도 더욱 편리하게 실 구매자들이 중요하게 여기는 제품의 특성들과 그에 대한 사용자 평가들을 손쉽게 확인할 수 있습니다. 아래의 예시들은 사용자에게 제공 가능한 서비스 시나리오 중 일부이며, 이외에도 서비스에 따라 다양한 방식으로 정보를 제공 할 수 있습니다. (블로그에 예시 나와있음)