

추천시스템

Index

1. 추천 시스템
- 개요
- 활용

2. 벡터
- 차원

3. 유사도 계산
- 유클리디언 거리
- 코사인 유사도

4. 추천 시스템 종류
- 지도 방식
- 클러스터링
- 콘텐츠 기반
- 협업 필터링
- 하이브리드

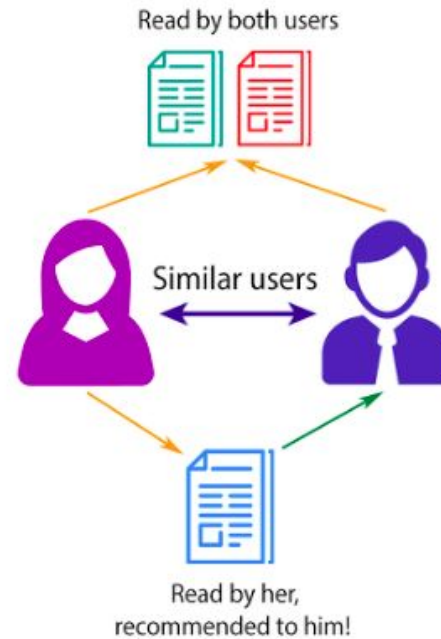
Recommendation

추천
시스템

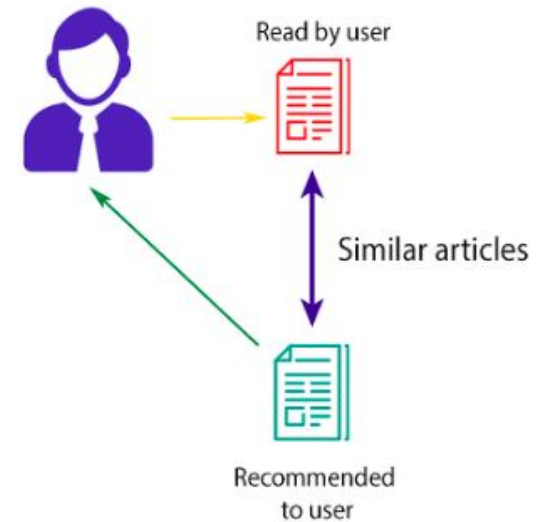
추천 시스템

- 추천시스템은 정보 필터링 (IF) 기술의 일종으로, 특정 사용자가 관심을 가질만한 정보를 추천하는 것

COLLABORATIVE FILTERING



CONTENT-BASED FILTERING



가치를 높이는 금융 인공지능 실무교육

Insightcampus

Recommendation

추천
시스템

영상 추천 - 유튜브

하루 종일 듣어도 좋은 노래 BEST 100곡
11:54:56

하루 종일 듣기 좋은 노래 🎵❤️ 일할 때 듣기 좋은 발라드 추천 광고없는 ...
주황색 ORANGE MUSIC
조회수 6.7만회 · 스트리밍 시간: 1개월 전

이 사버리!!
1:18:55

엘리베이터 치한 전치 12주 만들어버린 이다희! 김블유 처음부터 몰아보기
시네마로드
조회수 183만회 · 11개월 전

아무도.. "안 가요" 한국의 절반이 사라진다
9:14

"아무도 찾지 않아요"...한국의 절반이 사라진다 / SBS / 모아보는 뉴스
SBS 뉴스
조회수 44만회 · 5시간 전

김태리 연기 재미있다.. 2023년 꼭 봐야할 레전드 드라마가 될 김은희 작...
39:37

김시선
조회수 8만회 · 3시간 전

현장에서 난리 난 놀라운 용접기
6:00

누구나 쉽게 바로 용접이 가능한 특수한 기능의 논가스용접기! 용접도 이...
윤희들
조회수 34만회 · 4일 전

기자회견 도중 싸움이 났다!! UFC 역사상 가장 살벌했던 라이벌전
18:01

승률 100% 괴물 파이터를 갖고 노는 싸움의 신 ㄷㄷ #존존스
차도르
조회수 48만회 · 8일 전

미쳤다는 말밖에 안나온다 팀 전체를 바꿔버린 손흥민 이강인
6:04

둘만 템포가 다르네ㅋㅋ중계진도 감독도 경악시킨 손흥민과 이강인 수...
등현kldh
조회수 158만회 · 3일 전

진지하게 해야 되나 웃겨야 하나
13:58

[#유퀴즈온더블럭] 녹화 중에 웃길지 진지할지 아직도 선택 못 한 문상훈...
디글 :Diggle
조회수 68만회 · 3일 전

Recommendation

추천
시스템

영상 추천 - 유튜브

알 수 없는 유튜브 알고리즘이 나를 이곳으로 이끌었다.



하루 종일 듣어도 좋은 BEST 100 노래
11:54:56

하루 종일 듣기 좋은 노래 🎵❤️ 일할 때 듣기 좋은 발라드 추천 광고없는 ...
주황색 ORANGE MUSIC
조회수 6.7만회 · 스트리밍 시간: 1개월 전



이 서버로!!
1:18:55

엘레베이터 치한 전치 12주 만들어버린 이다희! 경복은 처음부터 몰아보기
조회수 183만회 · 1개월 전



아무도.. "안 가요" 한국의 절반이 사라진다
9:14

"아무도 찾지 않아요"...한국의 절반이 사라진다 / SBS / 모아보는 뉴스
조회수 1만회 · 3개월 전



김태리
39:37

김태리 연기 재미있다.. 2023년 꼭 봐야할 레전드 드라마가 될 김은희 작...
김시선
조회수 8만회 · 3시간 전



용접기 시대는 끝났다! 현장에서 난리 난 놀라운 용접기
6:00

누구나 쉽게 바로 용접이 가능한 특수한 기능의 논가스용접기! 용접도 이...
윤희들
조회수 34만회 · 4일 전



기자회견 도중 싸움이 났다!! UFC 역사상 가장 살벌했던 라이벌전
18:01

승률 100% 괴물 파이터를 갖고 노는 싸움의 신 ㄷㄷ #존존스
차도르
조회수 48만회 · 8일 전



미쳤다는 말밖에 안나온다 팀 전체를 바꿔버린 손흥민 이강인
6:04

둘만 템포가 다르네요ㅋㅋ중계진도 감독도 경악시킨 손흥민과 이강인 수...
등현kdh
조회수 158만회 · 3일 전



진지하게 해야 되나 웃겨야 하나
13:58

[#유퀴즈온더블럭] 녹화 중에 웃길지 진지할지 아직도 선택 못 한 문상훈...
디글 :Diggle
조회수 68만회 · 3일 전

Recommendation

추천
시스템

영상 추천 - 넷플릭스

오늘 대한민국의 TOP 10 시리즈



새로 올라온 콘텐츠



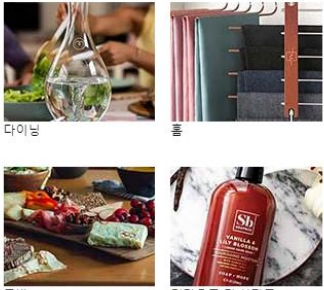







코미디 영화



Recommendation

추천
시스템

상품 추천 - 아마존

<p>내 공간 새 단장하기</p>  <p>다이닝 홀 주방 건강용품 및 화장품</p> <p>더 보기</p>	<p>AmazonBasics</p>  <p>더 보기</p>	<p>컴퓨터 및 액세서리</p>  <p>지금 쇼핑하세요</p>	<p>특가</p>  <p>지금 쇼핑하기</p>
<p>전자기기</p>  <p>더 보기</p>	<p>건강 및 퍼스널 케어</p>  <p>지금 쇼핑하세요</p>	<p>드레스</p>  <p>지금 쇼핑하세요</p>	<p>세련된 공간을 위한 간편한 업데이트</p>  <p>바로구니 하드웨어 포인트 가구 책지 및 패턴트</p> <p>홈 상품 쇼핑하기</p>

Recommendation

추천
시스템

음악 추천 - 애플 뮤직



Recommendation

추천
시스템

추천 시스템은 무엇을, 어떻게 추천할 수 있을까?

우리는 추천시스템을 크게 두 가지로 나눈 뒤, 하나하나씩 살펴봄으로써 추천시스템의 흐름을 배울 것이다.

- 무엇을

- 어떻게

Recommendation

추천
시스템

추천 시스템은 무엇을, 어떻게 추천할 수 있을까?

- 무엇을??

우선 무엇을 추천할지부터 생각해보자

Recommendation

추천
시스템

추천 시스템은 무엇을, 어떻게 추천할 수 있을까?

- 무엇을??

생각해보면, 우리가 추천받는 것들은
서비스 별로 모두 다른 종류의 것들이다.



책



영상

컴퓨터 및 액세서리



상품



노래

가치를 높이는 금융 인공지능 실무교육

Insightcampus


Recommendation

추천
시스템


추천 시스템은 무엇을, 어떻게 추천할 수 있을까?

- 무엇을??


하지만 추천하는 원리는 모두 같다. (아마도?)




초인류
김상균



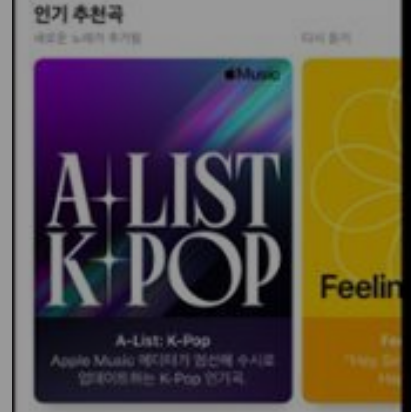
돈은 너로부터다
김종봉, 제갈현열



하루 종일 들어도 좋은 노래
주황색 ORANGE MUSIC
조희수 6.7만회 · 스트리밍 시간: 1개월 전



컴퓨터 및 액세서리



인기 추천곡
A+LIST K-POP
Apple Music
가치를 높이는 금융 인공지능 실무교육

책 영상 상품 노래

Recommendation

추천
시스템

추천 시스템은 무엇을, 어떻게 추천할 수 있을까?

- 무엇을??

하지만 추천하는 원리는 모두 같다. (아마도?)



Recommendation

추천
시스템

추천 시스템은 무엇을, 어떻게 추천할 수 있을까?

- 무엇을??

벡터화: 어떤 콘텐츠를 일련의 **숫자 목록**으로 표현하는 것

컨텐츠를 숫자로 표현하는
벡터화가 핵심

컨텐츠 = $\text{Vec}(0, 1, 2, 3, 4, 5, \dots, 789, 790)$

The collage shows four categories of content: 1. Books (책): '초인류' by Kim Sang-gun and '돈은 너로부터다' by Kim Jong-bok. 2. Videos (영상): A video player interface for '하루 종일 듣기 좋은 노래' (Songs to listen to all day) by ORANGE M.U. 3. Products (상품): A computer monitor and desk setup. 4. Music (노래): 'A+ LIST K+POP' and 'Feelin' albums. The text '컨텐츠 = Vec(0, 1, 2, 3, 4, 5, ..., 789, 790)' is overlaid on the collage, indicating that each item is represented by a vector of 791 numbers. Below the collage, the categories are labeled: 책, 영상, 상품, 노래. To the right of the '노래' label is the text '가치를 높이는 금융 인공지능 실무교육'.

책 영상 상품 노래 가치를 높이는 금융 인공지능 실무교육

Recommendation

추천
시스템

추천 시스템은 무엇을, 어떻게 추천할 수 있을까?

- 무엇을??

벡터화를 하면 무엇을 할 수 있을까?

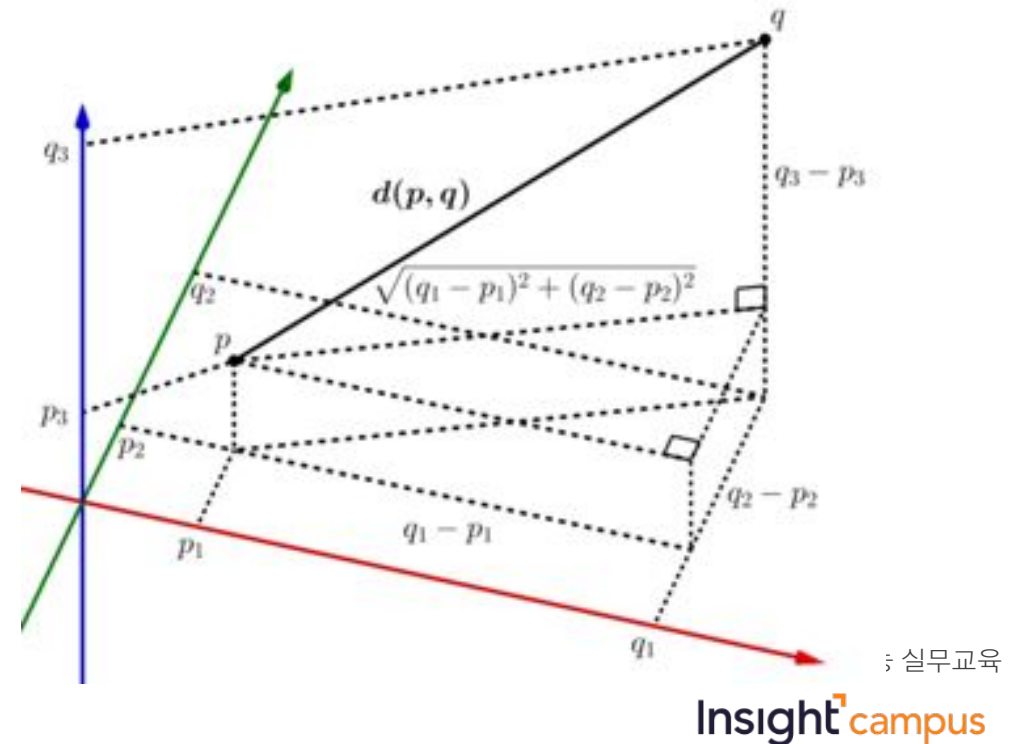
다양한 것이 가능하다.

1. 두 콘텐츠 사이의 **거리** 계산 (유클리디언 거리)
2. 두 콘텐츠 사이의 **각도** 계산 (코사인 유사도)

거리, 각도를 가지고,

두 콘텐츠가 서로 얼마나 가까운지 (=유사한지)
두 콘텐츠가 같은 방향을 바라보는지 (=유사한지)

를 가지고 콘텐츠를 추천할 수 있게 된다.



Recommendation

추천
시스템

추천 시스템은 무엇을, 어떻게 추천할 수 있을까?

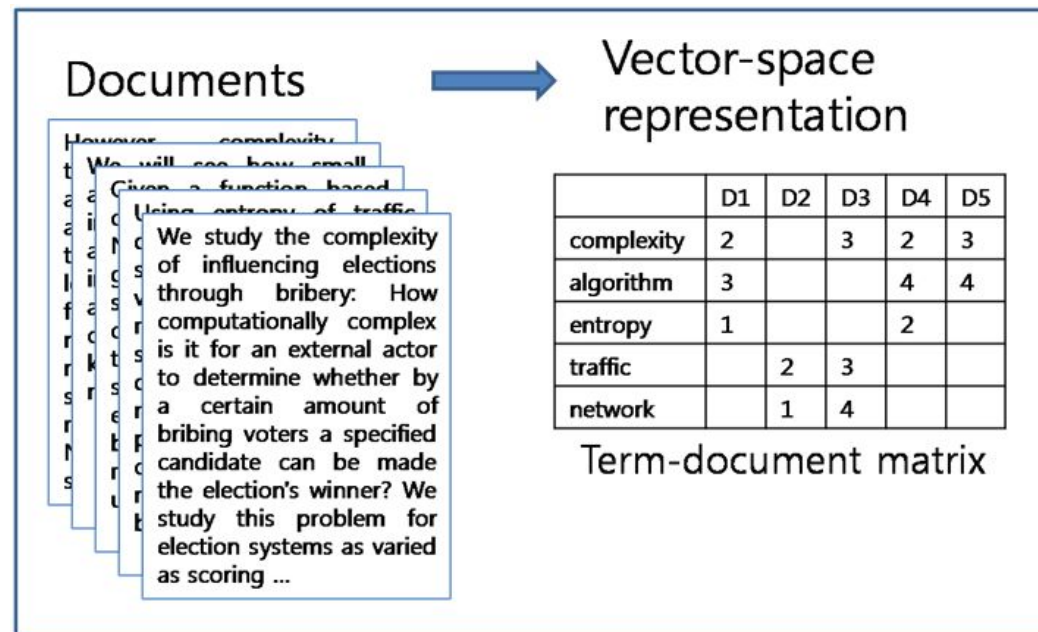
- 어떻게??

추천 시스템에서 컨텐츠(무엇을)를 추천하는데 핵심은 **벡터화**라는 것을 알았다.

그리고, 거리나 각도 계산을 통해 **유사도**를 판단할 수 있는 것도 알았다.

그럼 이 컨텐츠(영상, 책, 음악, 등 비정형 데이터)에 대해서 벡터화를 어떻게? 할 것인가가 자연스럽게 궁금해진다...

벡터화를 어떻게 할까..



가지를 높이는 금융 인공지능 실무교육

Recommendation

추천
시스템

추천 시스템은 무엇을, 어떻게 추천할 수 있을까?

- 어떻게??

일단 기본적으로 벡터화에는 어떤 정보든 활용할 수 있다. (성능은 그 다음 문제)

상품 정보

상품명 : 아이패드

브랜드 : Apple

세대 : 9

프로세서 : A13 바이오닉 칩

화면 : 10.2인치

용량 : 64GB

색깔 : 스페이스 그레이

가격 : 248.39 USD



Recommendation

추천
시스템

추천 시스템은 무엇을, 어떻게 추천할 수 있을까?

- 어떻게??

일단 기본적으로 벡터화에는 어떤 정보든 활용할 수 있다. (성능은 그 다음 문제)

상품 정보

상품명 : 아이패드

브랜드 : Apple

세대 : 9

프로세서 : A13 바이오닉 칩

화면 : 10.2인치

용량 : 64GB

색깔 : 스페이스 그레이

가격 : 248.39 USD



무교육

ous

Recommendation

추천
시스템

추천 시스템은 무엇을, 어떻게 추천할 수 있을까?

- 어떻게??

일단 기본적으로 벡터화에는 어떤 정보든 활용할 수 있다. (성능은 그 다음 문제)

상품 정보

상품명 : 아이패드

브랜드 : Apple

세대 : 9

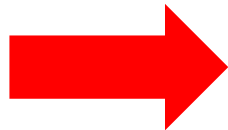
프로세서 : A13 바이오닉 칩

화면 : 10.2인치

용량 : 64GB

색깔 : 스페이스 그레이

가격 : 248.39 USD



상품 1 : (9.0, 10.2, 64.0,
248.39)



Recommendation

추천
시스템

추천 시스템은 무엇을, 어떻게 추천할 수 있을까?

- 어떻게??

일단 기본적으로 벡터화에는 어떤 정보든 활용할 수 있다. (성능은 그 다음 문제)

상품 정보

상품명 : 아이패드

브랜드 : Apple

세대 : 9

프로세서 : A13 바이오닉 칩

화면 : 10.2인치

용량 : 64GB

색깔 : 스페이스 그레이

가격 : 248.39 USD



(세대, 화면, 용량, 가격)

상품 1 : (9.0, 10.2, 64.0,
248.39)



Recommendation

추천
시스템

추천 시스템은 무엇을, 어떻게 추천할 수 있을까?

- 어떻게??

일단 기본적으로 벡터화에는 어떤 정보든 활용할 수 있다. (성능은 그 다음 문제)

상품 정보

상품명 : 아이패드

브랜드 : Apple

세대 : 9

프로세서 : A13 바이오닉 칩

화면 : 10.2인치

용량 : 64GB

색깔 : 스페이스 그레이

가격 : 248.39 USD



(세대, 화면, 용량, 가격) = 차원

상품 1 : (9.0, 10.2, 64.0,

248.39)

상품 2 : (7.0, 12.4, 256.0,

322.17)

⋮

상품 5 : (8.0, 8.7, 128.0,

189.39)



Recommendation

추천
시스템

추천 시스템은 무엇을, 어떻게 추천할 수 있을까?

- 어떻게??

이 벡터를 가지고 유사도를 계산해보자. (코사인 유사도)

상품 1 : (9.0, 10.2, 64.0,
248.39)
상품 2 : (7.0, 12.4, 256.0,
322.17)
⋮
상품 5 : (8.0, 8.7, 128.0,
189.39)



Recommendation

추천
시스템

추천 시스템은 무엇을, 어떻게 추천할 수 있을까?

- 어떻게??

이 벡터를 가지고 유사도를 계산해보자. (코사인 유사도)

상품 1 : (9.0, 10.2, 64.0,
248.39)
상품 2 : (7.0, 12.4, 256.0,
322.17)
⋮
상품 5 : (8.0, 8.7, 128.0,
189.39)

cos_sim	상품 1	상품 2	...	상품 5
상품 1	1	0.91334	...	0.94218
상품 2	0.91334	1	...	0.99684
...
상품 5	0.94218	0.99684	...	1



Recommendation

추천
시스템

추천 시스템은 무엇을, 어떻게 추천할 수 있을까?

- 어떻게??

구하긴 했는데 뭔가 아쉽다..

이 4개 특징(차원)만으로 상품을 잘 추천했다고 할 수 있을까?



상품 1 : (9.0, 10.2, 64.0,
248.39)

Recommendation

추천
시스템

추천 시스템은 무엇을, 어떻게 추천할 수 있을까?

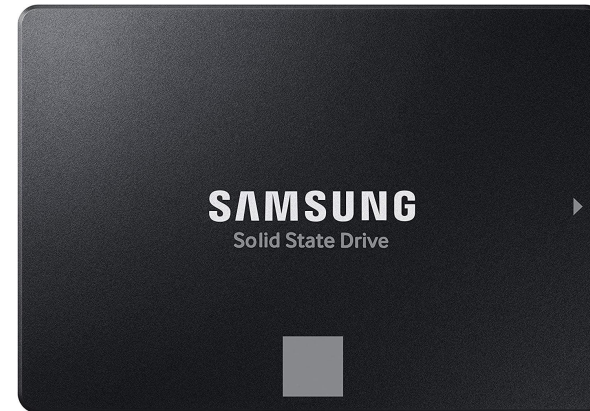
- 어떻게??

구하긴 했는데 뭔가 아쉽다..

이 4개 특징(차원)만으로 상품을 잘 추천했다고 할 수 있을까?
그리고, 아이패드가 아닌 상품은?



상품 1 : (9.0, 10.2, 64.0,
248.39)



상품 2 : (870.0, 2.5, 1024, 47.0)

가치를 높이는 금융 인공지능 실무교육

Insightcampus

Recommendation

추천
시스템

추천 시스템은 무엇을, 어떻게 추천할 수 있을까?

- 어떻게??

구하긴 했는데 뭔가 아쉽다..

두 상품의 차원 수=4로 같지만, 각 차원이 같은 내용이라고 할 수 있을까? 그리고 만약 차원수가 다르다면?
(코사인 유사도는 기본적으로 같은 차원수의 벡터끼리만 계산 할 수 있다.)

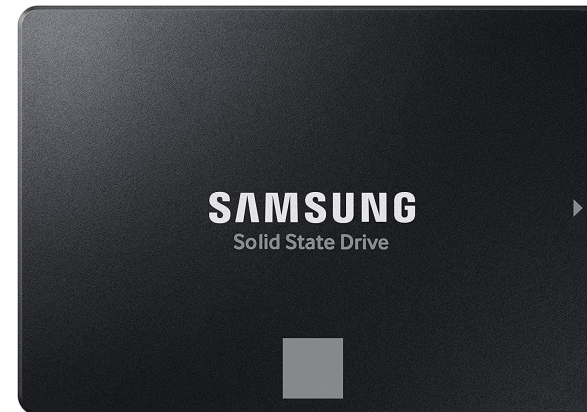
(세대, 화면, 용량, 가격)



상품 1 : (9.0, 10.2, 64.0,
248.20)

(버전, 크기, 용량, 가격)

!=



상품 2 : (870.0, 2.5, 1024, 47.0)

치를 높이는 금융 인공지능 실무교육

Insightcampus

Recommendation

추천
시스템

추천 시스템은 무엇을, 어떻게 추천할 수 있을까?

- 어떻게??

벡터화에는 숫자(정형 데이터)만 활용할 수 있을까?

상품 정보

상품명 : 아이패드

브랜드 : Apple

세대 : 9

프로세서 : A13 바이오닉 칩

화면 : 10.2인치

용량 : 64GB

색깔 : 스페이스 그레이

가격 : 248.39 USD



Recommendation

추천
시스템

추천 시스템은 무엇을, 어떻게 추천할 수 있을까?

- 어떻게??

벡터화에는 숫자(정형 데이터)만 활용할 수 있을까?
텍스트 데이터는? 영상은? 이미지는? 음악은?

상품 정보

상품명 : 아이패드

브랜드 : Apple

세대 : 9

프로세서 : A13 바이오닉 칩

화면 : 10.2인치

용량 : 64GB

색깔 : 스페이스 그레이

가격 : 248.39 USD



Recommendation

추천
시스템

추천 시스템은 무엇을, 어떻게 추천할 수 있을까?

- 어떻게??

텍스트 데이터, 영상 데이터, 음성 데이터 등 모두 숫자로 표현하려는 시도가 있다. (궁금하면 한번 찾아보자.)

상품 정보

상품명 : 아이패드

브랜드 : Apple

세대 : 9

프로세서 : A13 바이오닉 칩

화면 : 10.2인치

용량 : 64GB

색깔 : 스페이스 그레이

가격 : 248.39 USD



Recommendation

추천
시스템

추천 시스템은 무엇을, 어떻게 추천할 수 있을까?

- 어떻게??

이러한 데이터 중 범주형(카테고리) 데이터는 숫자로 표현하기가 그나마 수월하다.

상품 정보

상품명 : 아이패드 (아이패드 = 0, SSD = 1)

브랜드 : Apple (Apple = 0, 삼성 = 1)

세대 : 9

프로세서 : A13 바이오닉 칩 (Apple = 0, Intel = 1)

화면 : 10.2인치

용량 : 64GB

색깔 : 스페이스 그레이 (색깔 별로 0~9)

가격 : 248.39 USD



Recommendation

추천
시스템

추천 시스템은 무엇을, 어떻게 추천할 수 있을까?

- 어떻게??

이렇게 상품별로 공통적인 속성을 뽑아서 상품별 같은 차원 수의 벡터를 만들면, 서로 다른 종류여도, 상품별 유사도를 계산하여 추천할 수 있게 된다.

텍스트를 벡터화하는 방법 중 하나인 단어 - 문서 행렬

-	cat	dog	I	like
문서1	0	1	1	1
문서2	1	0	1	1
문서3	2	0	2	2



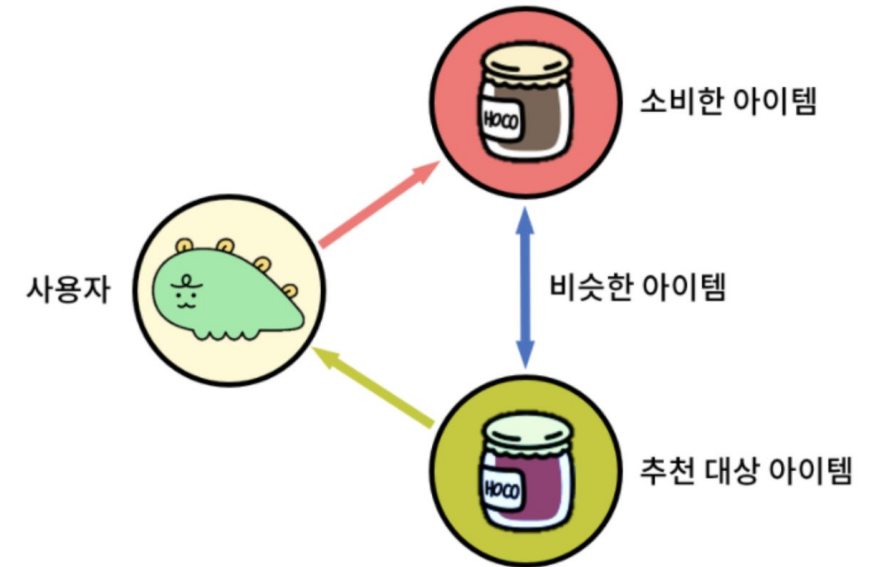
Recommendation

추천
시스템

전체적으로 정리해보자.

추천 시스템은

1. 상품, 콘텐츠를 **벡터화(=숫자로 표현)**하여
2. 벡터 간 **유사도**를 계산한 다음,
3. 유사도가 높은 순서대로 정렬하여, **유사도가 높은 순서대로 상품을 추천**한다



가치를 높이는 금융 인공지능 실무교육

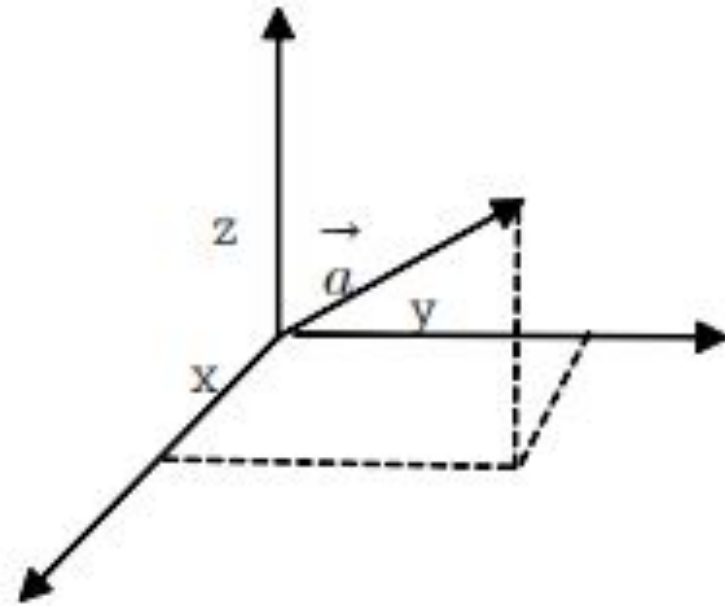
Insightcampus

Vector

벡
터

벡터

- 수학개념으로 크기와 방향을 갖는 물리량
- 이론적으로는 그렇지만, 머신러닝 & 딥러닝에서 벡터라고 하는 건 주로 특징(차원)들로 이루어진 숫자들이라고 이해하는게 편하다.
- 이 숫자들은 직선에서(1차원), 평면상에서(2차원), 혹은 공간상(3차원 이상)에서 특정한 위치에 점으로써 표현된다.

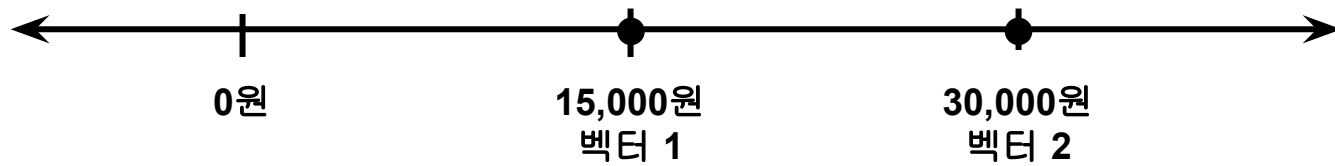


Vector

벡터

벡터의 예

- 가격이 15000원, 30000원인 벡터 (1차원 직선)



가치를 높이는 금융 인공지능 실무교육

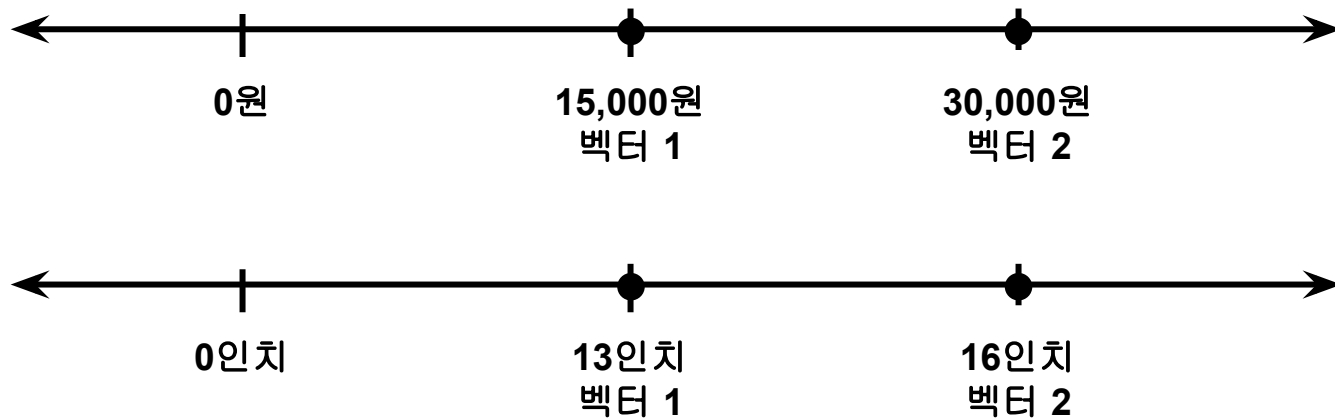
Insightcampus

Vector

벡터

벡터의 예

- 가격이 15000원, 30000원인 벡터 (1차원 직선)
- 여기에 화면 크기 차원을 추가 (1차원 직선)

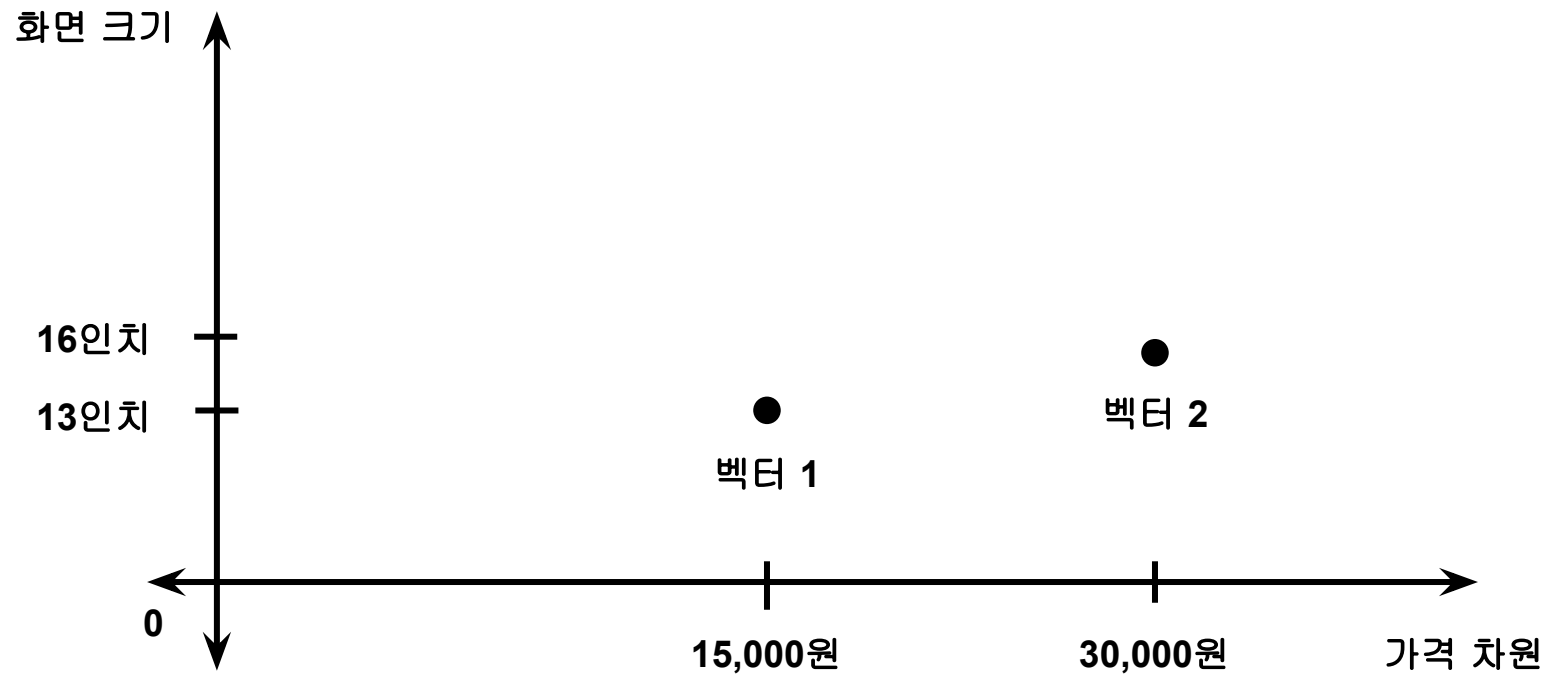


Vector

벡터

벡터의 예

- 가격, 화면 크기 (15000, 13), (30000, 16)의 두 개 벡터를 2차원 공간에 표현



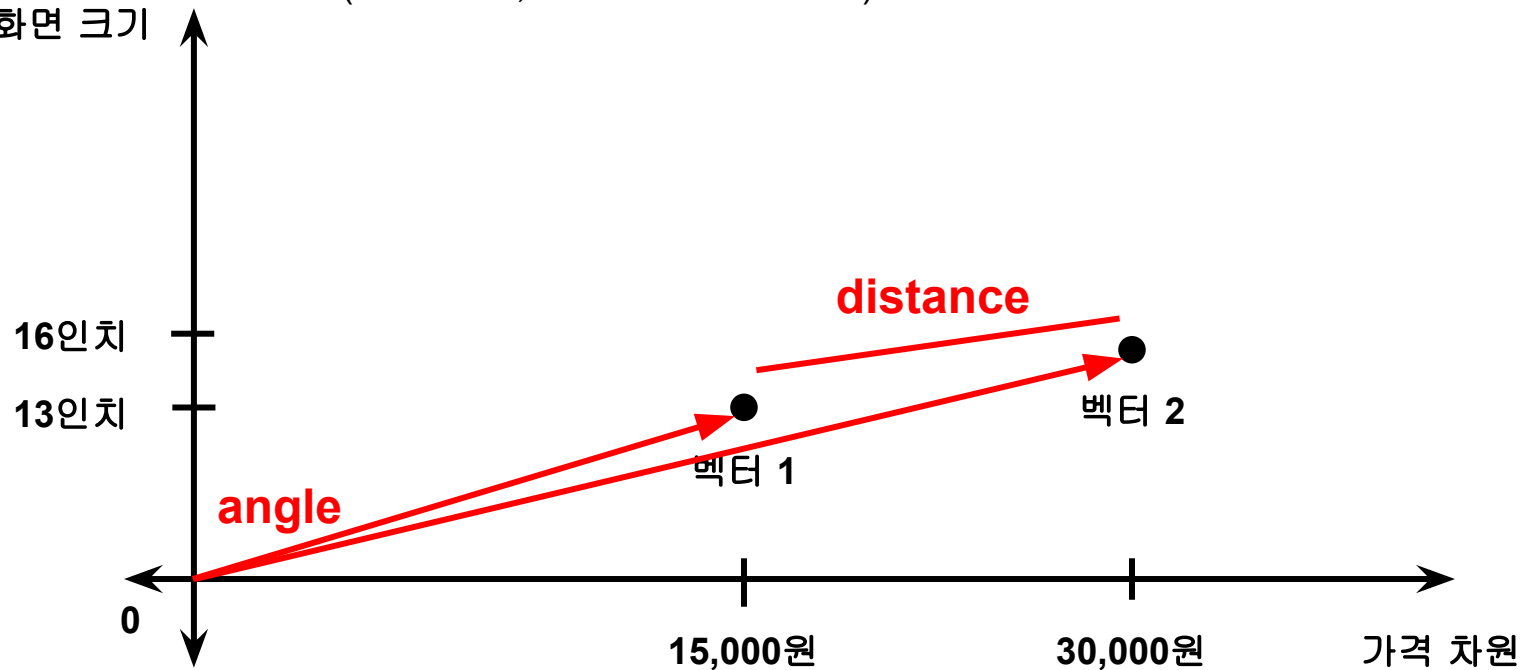
Vector

벡터

벡터의 예

- 차원(특징) 수가 늘어나면 늘어날수록 사람이 볼 수 있는 방법으로 표현하기는 어렵다. 하지만 개념적으로 공간상에 있는 특정한 점이라고 생각하면 그 점들이 거리가 가까운지, 먼지원점에서 두 점이 같은 방향에 있는지 계산할 수 있다. (거리 계산, 코사인 유사도 계산)

화면 크기

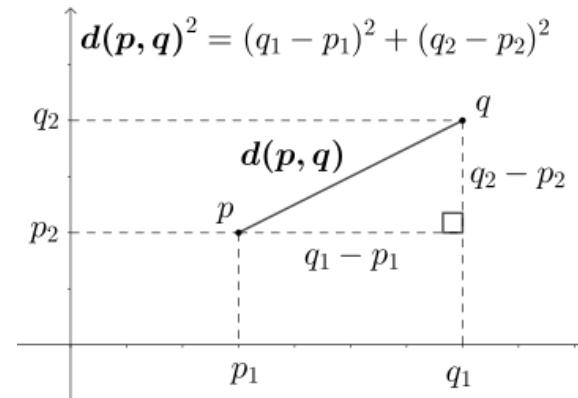


Similarity

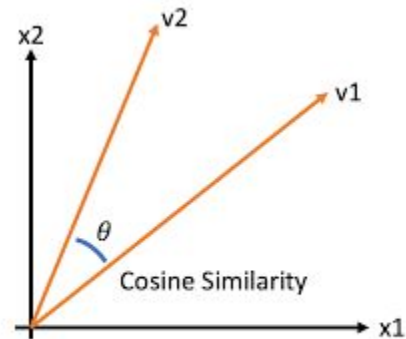
유사
도

유사도 계산

- 유클리디언 거리



- 코사인 유사도



Similarity

유사
도

유사도 계산

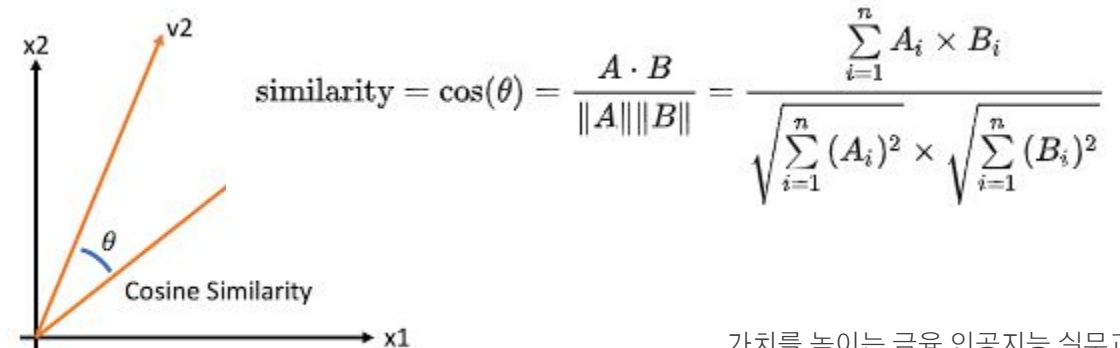
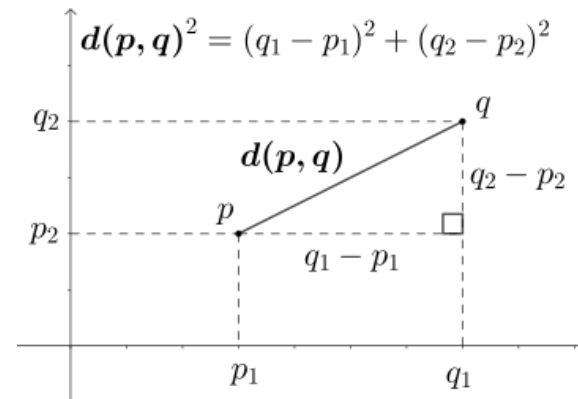
- 유클리디언 거리

```
import numpy as np
a = np.array((1, 2, 3))
b = np.array((4, 5, 6))
dist = np.sqrt(np.sum(np.square(a-b))) print(dist)
```

- 코사인 유사도

```
from numpy import dot
from numpy.linalg import norm
import numpy as np
```

```
def cos_sim(A, B):
    return dot(A, B)/(norm(A)*norm(B))
```



가치를 높이는 금융 인공지능 실무교육

Insightcampus

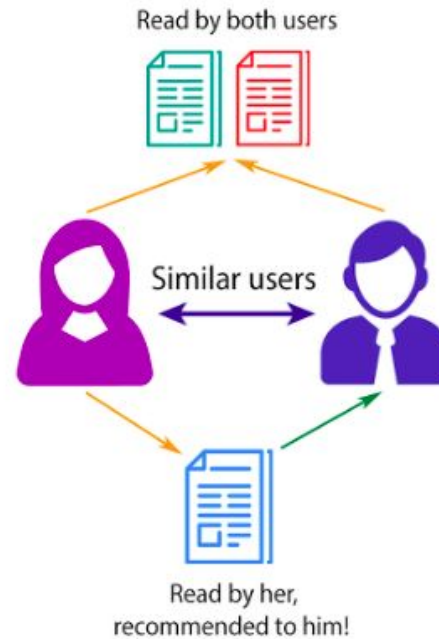
Recommendation System

추천 시스템
종류

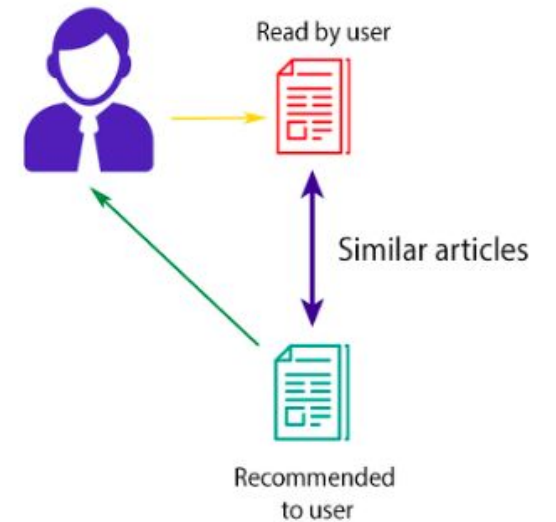
추천시스템 종류

- 지도 학습법 (target 여부 O)
- 클러스터링 (군집화)
- 콘텐츠 기반 필터링
- 협업 필터링
- 하이브리드 필터링

COLLABORATIVE FILTERING



CONTENT-BASED FILTERING



가치를 높이는 금융 인공지능 실무교육

Insight^{campus}

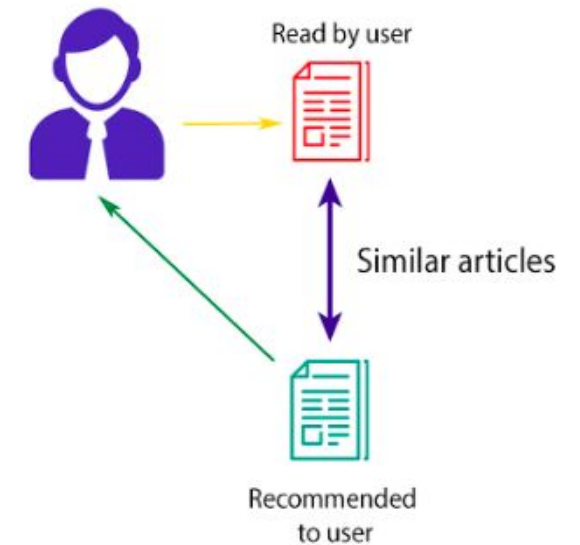
Recommendation System

추천 시스템
종류

컨텐츠 기반 필터링

- 사용자가 특정한 아이템을 선호하는 경우, 그 아이템과 비슷한 콘텐츠를 가진 다른 아이템을 추천하는 방식

CONTENT-BASED FILTERING



가치를 높이는 금융 인공지능 실무교육

Insight^{campus}

Recommendation System

추천 시스템
종류

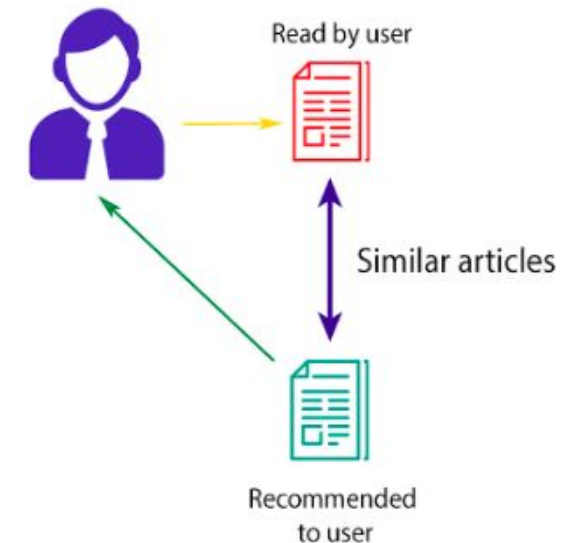
컨텐츠 기반 필터링

- 사용자가 특정한 아이템을 선호하는 경우, 그 아이템과 비슷한 콘텐츠를 가진 다른 아이템을 추천하는 방식

기본적으로 추천시스템의 원리는 벡터화, 유사도 계산이라고 했다.

컨텐츠 기반 필터링에서 **벡터**는 무엇이 될 수 있을까?

CONTENT-BASED FILTERING



가치를 높이는 금융 인공지능 실무교육

Insightcampus

Recommendation System

추천 시스템
종류

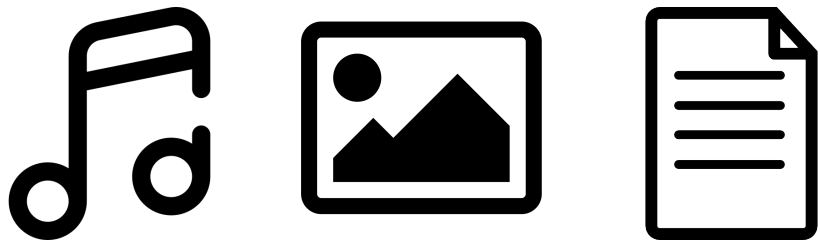
컨텐츠 기반 필터링

- 사용자가 특정한 아이템을 선호하는 경우, 그 아이템과 비슷한 콘텐츠를 가진 다른 아이템을 추천하는 방식

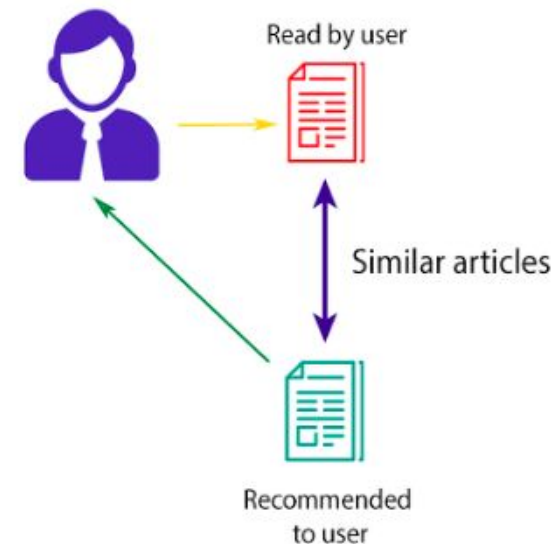
기본적으로 추천시스템의 원리는 벡터화, 유사도 계산이라고 했다.

컨텐츠 기반 필터링에서 **벡터**는 무엇이 될 수 있을까?

이미지, 음성, 태그, 상품 설명서, 리뷰 등과 같은 **컨텐츠에 대한 설명**으로부터 콘텐츠의 특징 벡터를 추출할 수 있다.



CONTENT-BASED FILTERING



가치를 높이는 금융 인공지능 실무교육

Insightcampus

Recommendation System

추천 시스템
종류

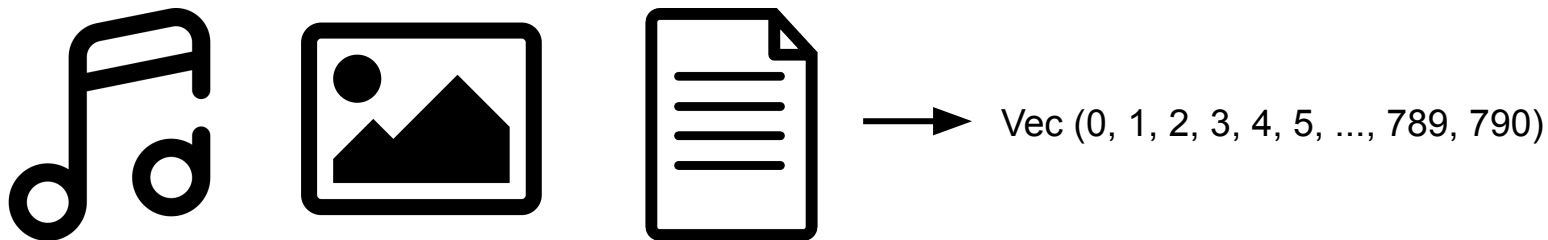
컨텐츠 기반 필터링

- 사용자가 특정한 아이템을 선호하는 경우, 그 아이템과 비슷한 콘텐츠를 가진 다른 아이템을 추천하는 방식

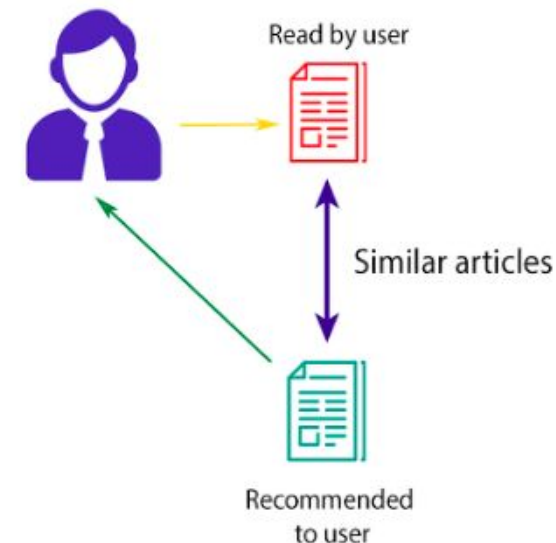
기본적으로 추천시스템의 원리는 벡터화, 유사도 계산이라고 했다.

컨텐츠 기반 필터링에서 **벡터**는 무엇이 될 수 있을까?

이미지, 음성, 태그, 상품 설명서, 리뷰 등과 같은 **컨텐츠에 대한 설명**으로부터 콘텐츠의 특징 벡터를 추출할 수 있다.



CONTENT-BASED FILTERING



가치를 높이는 금융 인공지능 실무교육

Insightcampus

Recommendation System

추천 시스템
종류

컨텐츠 기반 필터링

- 사용자가 특정한 아이템을 선호하는 경우, 그 아이템과 비슷한 콘텐츠를 가진 다른 아이템을 추천하는 방식

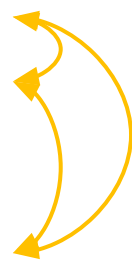
유사도 계산은 쉽다. 콘텐츠별로 만들어진 벡터를 가지고, 각 벡터끼리 유사도 계산 함수를 적용하면 된다.

Vec1 = (0, 1, 2, 3, 4, 5, ..., 789, 790)

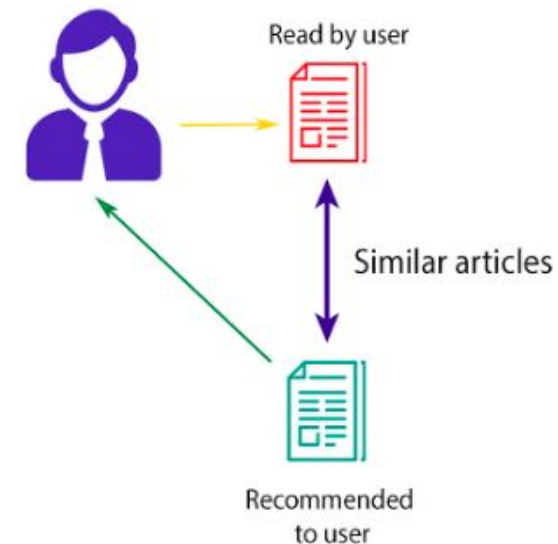
Vec2 = (0, 1, 2, 3, 4, 5, ..., 789, 790)

⋮

Vec5 = (0, 1, 2, 3, 4, 5, ..., 789, 790)



CONTENT-BASED FILTERING



가치를 높이는 금융 인공지능 실무교육

Insightcampus

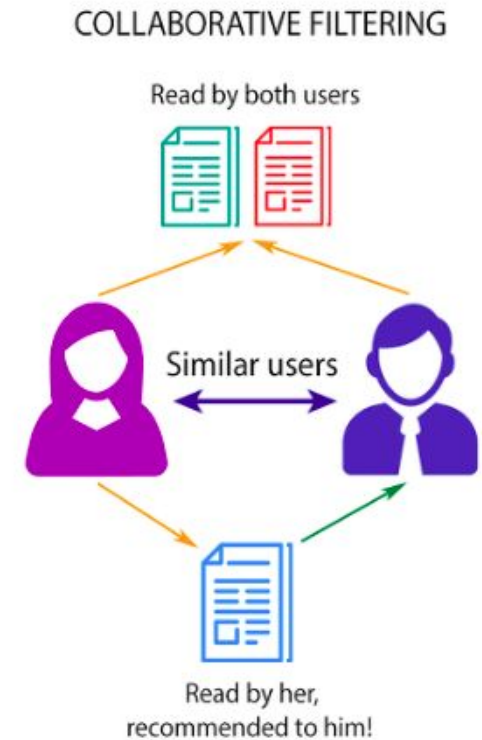
Recommendation System

추천 시스템
종류

협업 필터링

- 사용자와 아이템 간의 상호 상관 관계를 분석하여 새로운 사용자-아이템 관계를 찾아주는 것으로 사용자의 과거 경험과 행동 방식(**User Behavior**)에 의존하여 추천하는 시스템

협업 필터링에서 사용하는 **벡터**는 무엇일까?



가치를 높이는 금융 인공지능 실무교육

Insightcampus

Recommendation System

추천 시스템

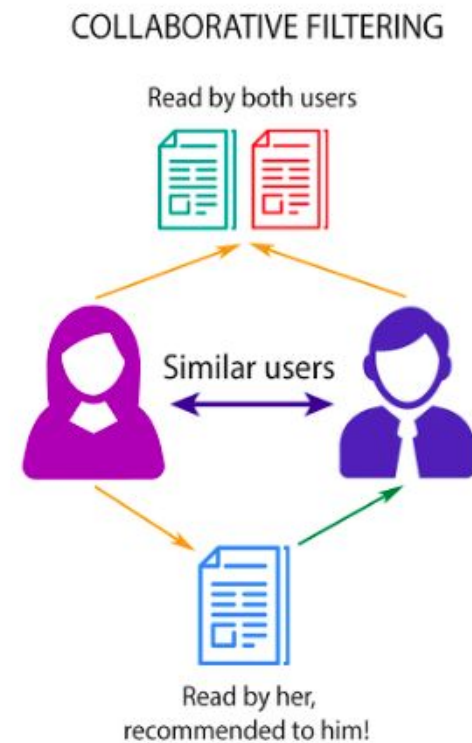
종류

협업 필터링

- 사용자와 아이템 간의 상호 상관 관계를 분석하여 새로운 사용자-아이템 관계를 찾아주는 것으로 사용자의 과거 경험과 행동 방식(**User Behavior**)에 의존하여 추천하는 시스템

협업 필터링에서 사용하는 **벡터**는 무엇일까?

사용자 - 아이템 간의 관계를 분석하는 것이라고 했으니, **둘의 관계**에 대한 정보를 담고 있는 벡터일 것이다. 무엇이 있을까...?



가치를 높이는 금융 인공지능 실무교육

Recommendation System

추천 시스템
종류

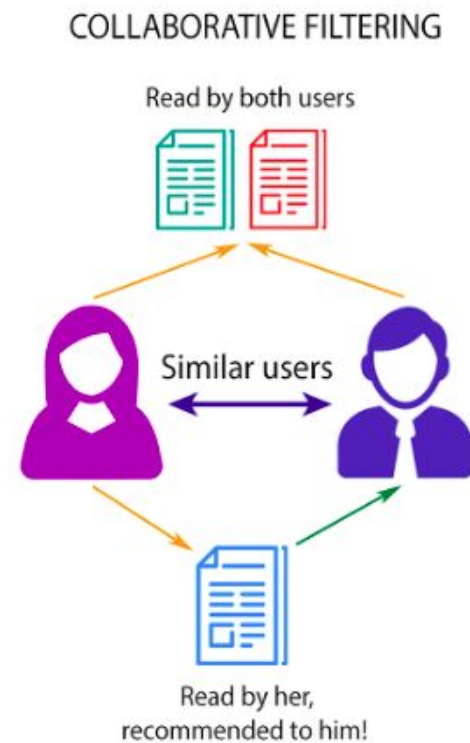
협업 필터링

- 사용자와 아이템 간의 상호 상관 관계를 분석하여 새로운 사용자-아이템 관계를 찾아주는 것으로 사용자의 과거 경험과 행동 방식(**User Behavior**)에 의존하여 추천하는 시스템

협업 필터링에서 사용하는 **벡터**는 무엇일까?

사용자 - 아이템 간의 관계를 분석하는 것이라고 했으니, **둘의 관계**에 대한 정보를 담고 있는 벡터일 것이다. 무엇이 있을까...?

그렇다. 사용자가 남긴 아이템에 대한 **평가(리뷰)**다.



가치를 높이는 금융 인공지능 실무교육

Recommendation System

추천 시스템
종류

협업 필터링

- 사용자와 아이템 간의 상호 상관 관계를 분석하여 새로운 사용자-아이템 관계를 찾아주는 것으로 사용자의 과거 경험과 행동 방식(**User Behavior**)에 의존하여 추천하는 시스템

협업 필터링에서 사용하는 **벡터**는 무엇일까?

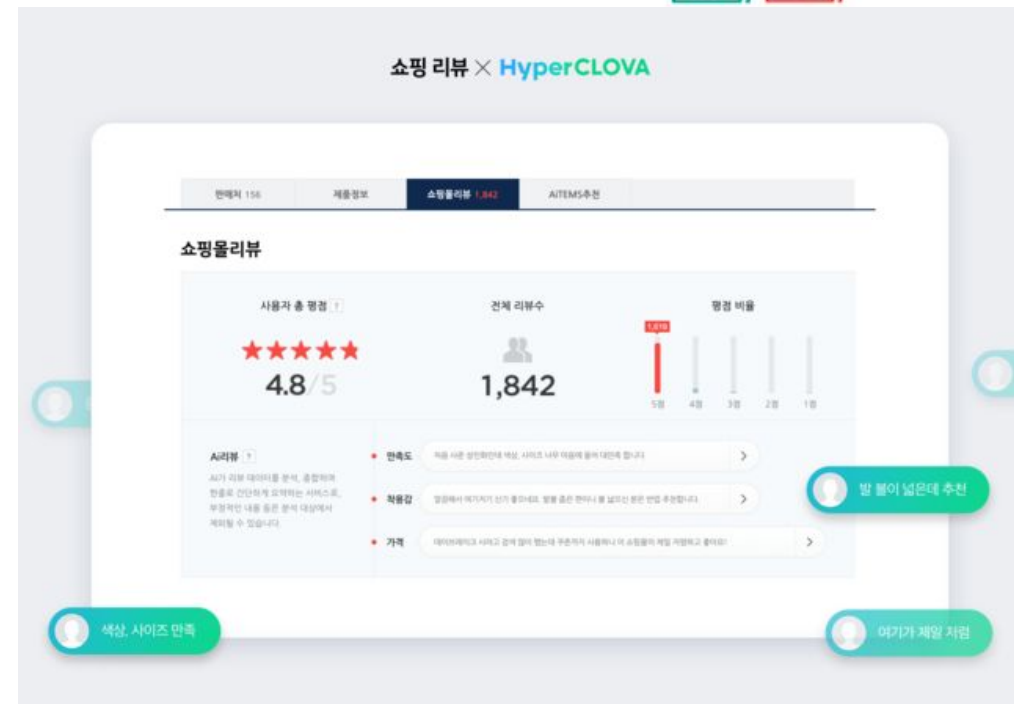
사용자 - 아이템 간의 관계를 분석하는 것이라고 했으니, **둘의 관계**에 대한 정보를 담고 있는 벡터일 것이다. 무엇이 있을까...?

그렇다. 사용자가 남긴 아이템에 대한 **평가(리뷰)**다.

리뷰 데이터는 사용자가 느낀 해당 아이템의 선호 정도를 나타낸다.

COLLABORATIVE FILTERING

Read by both users



Recommendation System

추천 시스템

종류

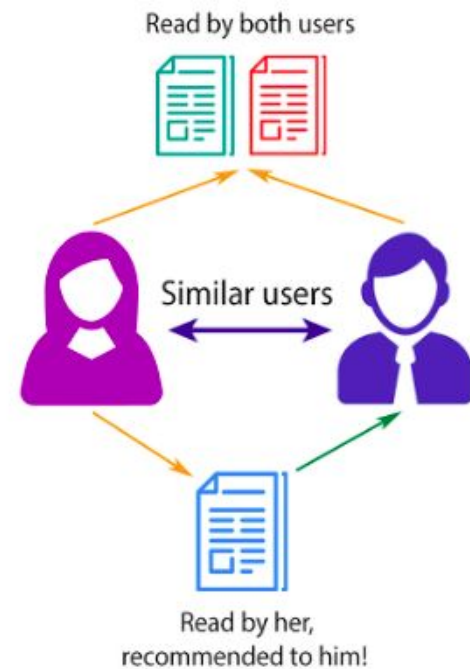
협업 필터링

- 사용자와 아이템 간의 상호 상관 관계를 분석하여 새로운 사용자-아이템 관계를 찾아주는 것으로 사용자의 과거 경험과 행동 방식(**User Behavior**)에 의존하여 추천하는 시스템

리뷰 데이터

	userId	movieId	rating	timestamp	title	genres
0	1	1	4.0	964982703	Toy Story (1995)	Adventure Animation Children Comedy Fantasy
1	5	1	4.0	847434962	Toy Story (1995)	Adventure Animation Children Comedy Fantasy

COLLABORATIVE FILTERING



가치를 높이는 금융 인공지능 실무교육

Insightcampus

Recommendation System

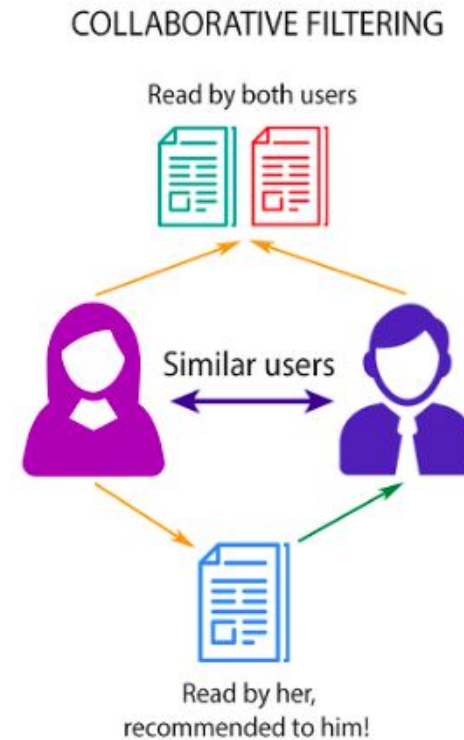
추천 시스템
종류

협업 필터링

- 사용자와 아이템 간의 상호 상관 관계를 분석하여 새로운 사용자-아이템 관계를 찾아주는 것으로 사용자의 과거 경험과 행동 방식(**User Behavior**)에 의존하여 추천하는 시스템

리뷰 행렬 (=matrix, vector를 행으로 쌓은 형태)

	items				
user1	0	4	0	3	2
user2	0	5	0	1	1
user3	0	0	3	4	2
user4	1	4	2	0	0
user5	2	4	1	3	2



가치를 높이는 금융 인공지능 실무교육

Insightcampus

Recommendation System

추천 시스템
종류

협업 필터링

- 사용자와 아이템 간의 상호 상관 관계를 분석하여 새로운 사용자-아이템 관계를 찾아주는 것으로 사용자의 과거 경험과 행동 방식(**User Behavior**)에 의존하여 추천하는 시스템

이 벡터들을 통해, user간 유사도를 계산할 수 있다.

	items				
user1	0	4	0	3	2
user2	0	5	0	1	1
user3	0	0	3	4	2
user4	1	4	2	0	0
user5	2	4	1	3	2



가치를 높이는 금융 인공지능 실무교육

Recommendation System

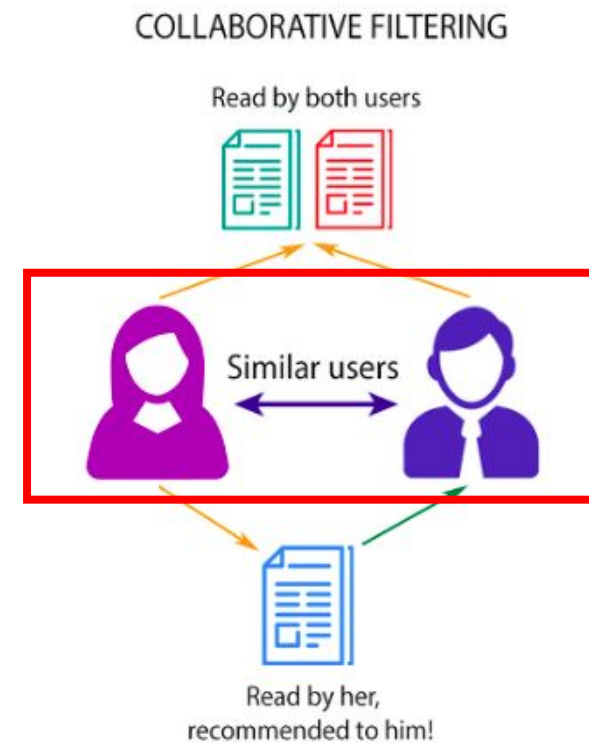
추천 시스템
종류

협업 필터링

- 사용자와 아이템 간의 상호 상관 관계를 분석하여 새로운 사용자-아이템 관계를 찾아주는 것으로 사용자의 과거 경험과 행동 방식(**User Behavior**)에 의존하여 추천하는 시스템

이 벡터들을 통해, **user**간 유사도를 계산할 수 있다.
(=나와 유사한 유저를 확인할 수 있다.)

	items				
user1	0	4	0	3	2
user2	0	5	0	1	1
user3	0	0	3	4	2
user4	1	4	2	0	0
user5	2	4	1	3	2



가치를 높이는 금융 인공지능 실무교육

Insightcampus

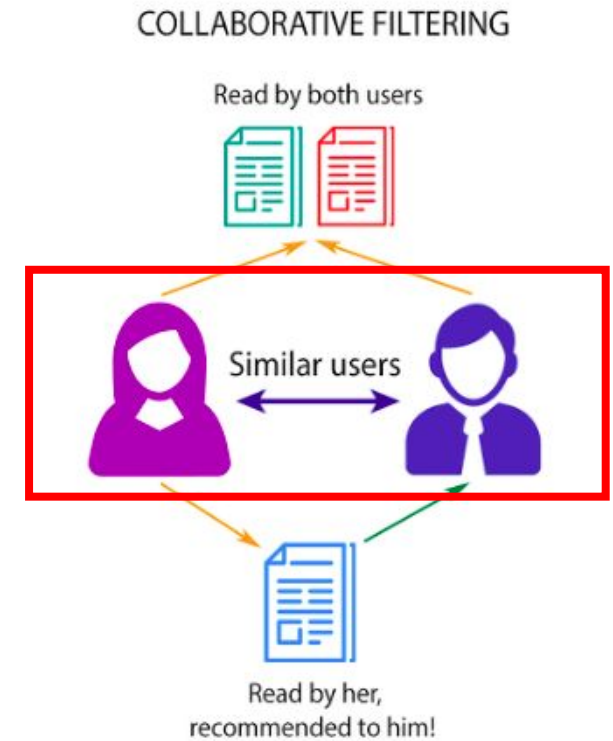
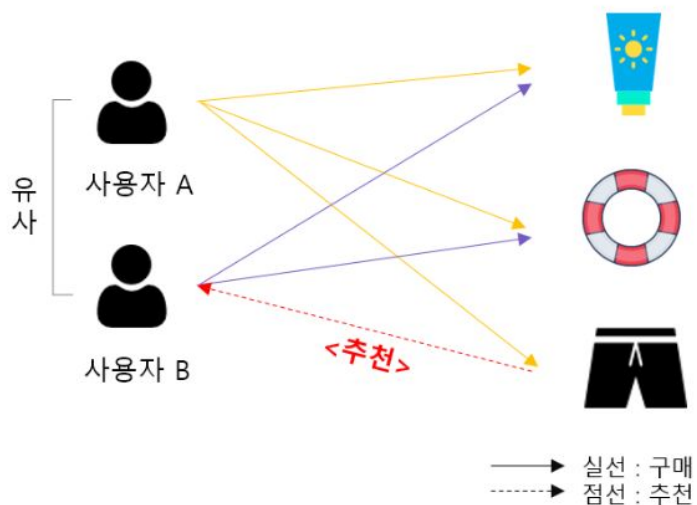
Recommendation System

추천 시스템
종류

협업 필터링

- 사용자와 아이템 간의 상호 상관 관계를 분석하여 새로운 사용자-아이템 관계를 찾아주는 것으로 사용자의 과거 경험과 행동 방식(**User Behavior**)에 의존하여 추천하는 시스템

이 벡터들을 통해, **user**간 유사도를 계산할 수 있다.
(=나와 유사한 유저를 확인할 수 있다.)



가치를 높이는 금융 인공지능 실무교육

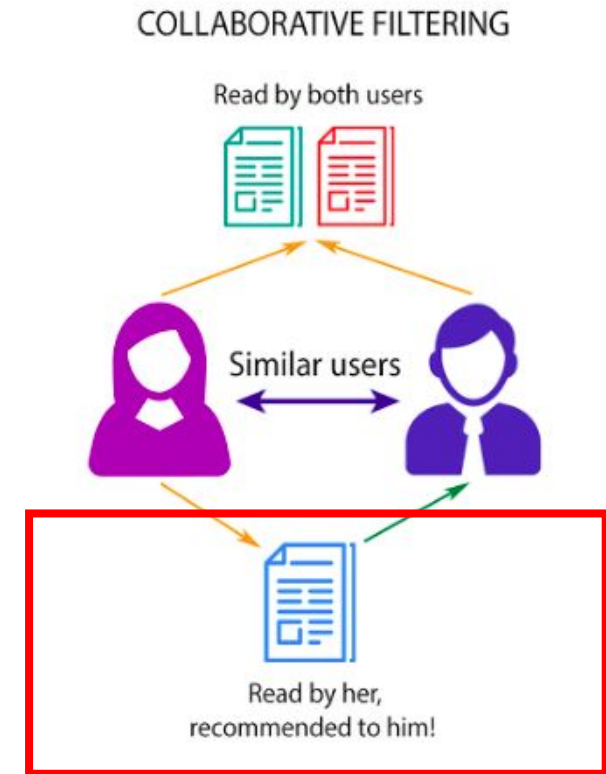
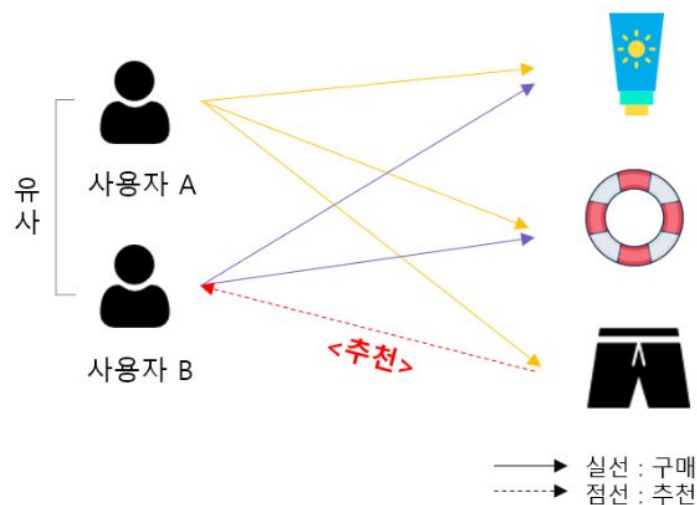
Recommendation System

추천 시스템
종류

협업 필터링

- 사용자와 아이템 간의 상호 상관 관계를 분석하여 새로운 사용자-아이템 관계를 찾아주는 것으로 사용자의 과거 경험과 행동 방식(User Behavior)에 의존하여 추천하는 시스템

이렇게 사용자 A와 사용자 B가 유사하다고 할 때, 구매하지 않은 제품이 있다면, 해당 상품을 추천해주는 것이 **협업 필터링**이다.



가치를 높이는 금융 인공지능 실무교육

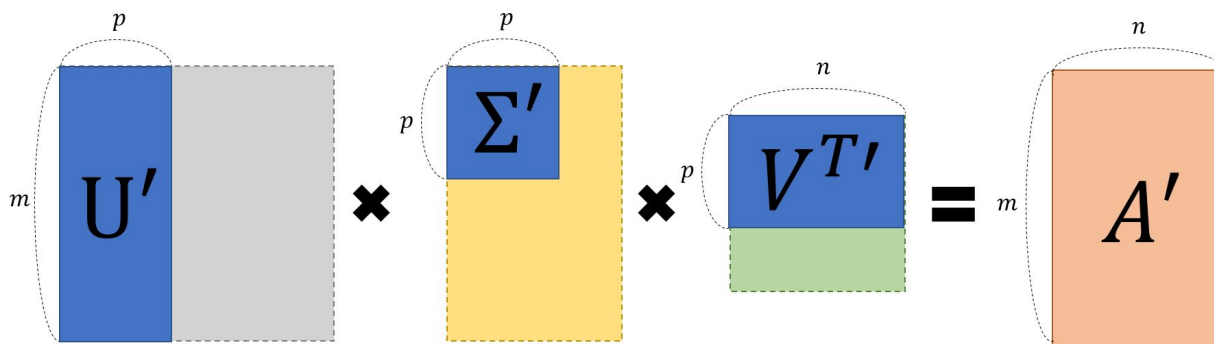
Recommendation System

추천 시스템
종류

협업 필터링

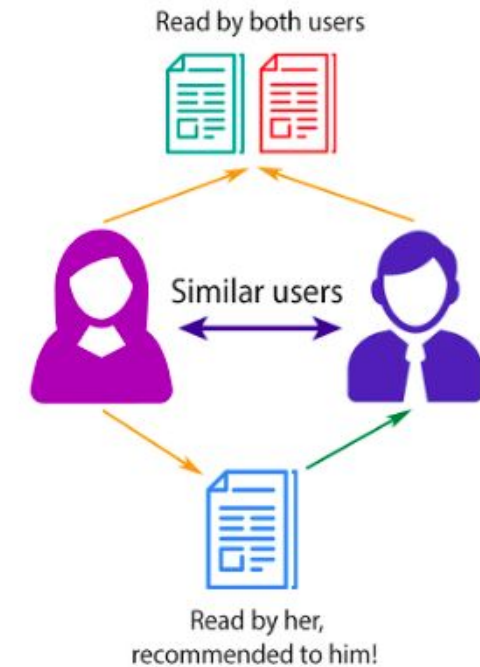
- 사용자와 아이템 간의 상호 상관 관계를 분석하여 새로운 사용자-아이템 관계를 찾아주는 것으로 사용자의 과거 경험과 행동 방식(User Behavior)에 의존하여 추천하는 시스템

잠재의미분석 SVD (유저에 대한 잠재 벡터, 아이템에 대한 잠재 벡터 생성)



The diagram illustrates the SVD decomposition of a matrix A' into three matrices: U' , Σ' , and $V^{T'}$. Matrix U' is of size $m \times p$ (blue and grey). Matrix Σ' is of size $p \times p$ (yellow with a blue top-left corner). Matrix $V^{T'}$ is of size $p \times n$ (blue and green). The equation is $U' \times \Sigma' \times V^{T'} = A'$, where A' is of size $m \times n$ (orange). Dimensions m , p , and n are indicated by dashed lines and labels.

COLLABORATIVE FILTERING



Recommendation System

추천 시스템
종류

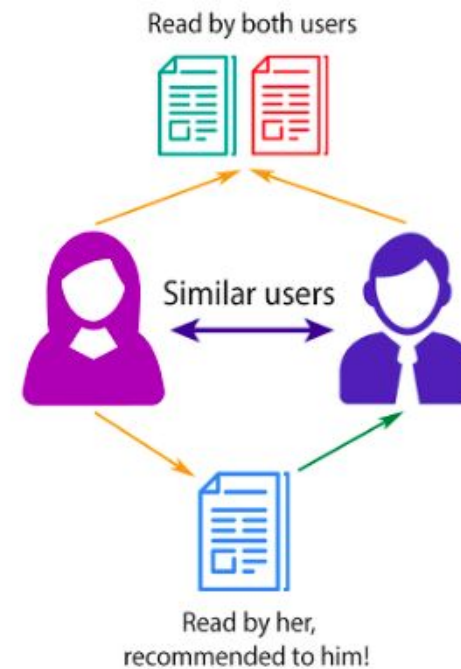
협업 필터링

- 사용자와 아이템 간의 상호 상관 관계를 분석하여 새로운 사용자-아이템 관계를 찾아주는 것으로 사용자의 과거 경험과 행동 방식(**User Behavior**)에 의존하여 추천하는 시스템

또한, 이 리뷰 행렬은 다르게도 활용해볼 수 있는데,

	items				
user1	0	4	0	3	2
user2	0	5	0	1	1
user3	0	0	3	4	2
user4	1	4	2	0	0
user5	2	4	1	3	2

COLLABORATIVE FILTERING



가치를 높이는 금융 인공지능 실무교육

Insightcampus

Recommendation System

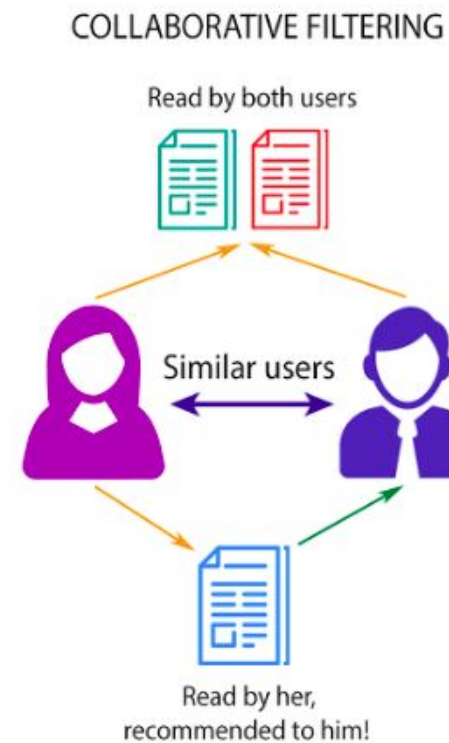
추천 시스템
종류

협업 필터링

- 사용자와 아이템 간의 상호 상관 관계를 분석하여 새로운 사용자-아이템 관계를 찾아주는 것으로 사용자의 과거 경험과 행동 방식(**User Behavior**)에 의존하여 추천하는 시스템

또한, 이 리뷰 행렬은 다르게도 활용해볼 수 있는데, 리뷰 행렬을 전치(**Transpose**)함으로써, **Item**에 대한 **vector**를 구할 수 있고, 이를 통해 콘텐츠 기반 필터링도 적용할 수 있다.

	users				
item1	0	0	0	1	2
item2	4	5	0	4	4
item3	0	0	3	2	1
item4	3	1	4	0	3
item5	2	1	2	0	2



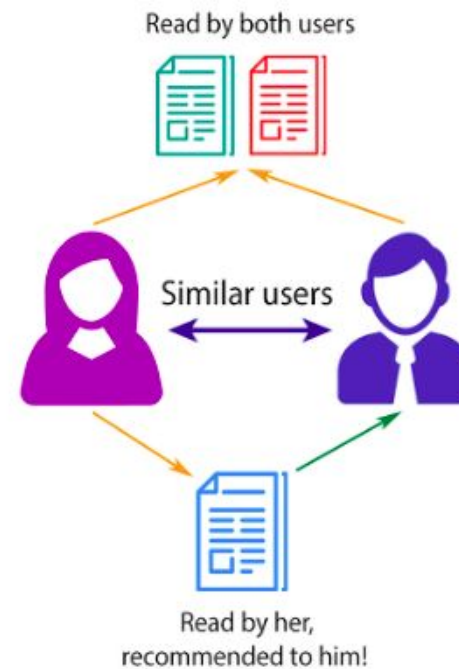
Recommendation System

추천 시스템
종류

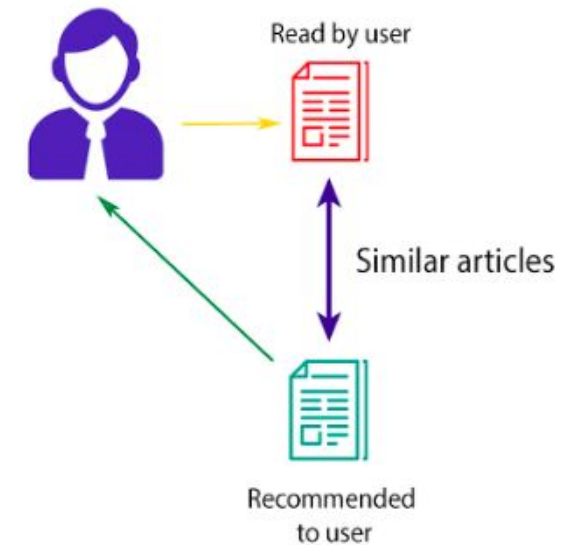
하이브리드 필터링

- 앞에서 다루었던 콘텐츠 기반 필터링과 협업 필터링을 다시
생각해보자.

COLLABORATIVE FILTERING



CONTENT-BASED FILTERING



교육

Recommendation System

추천 시스템
종류

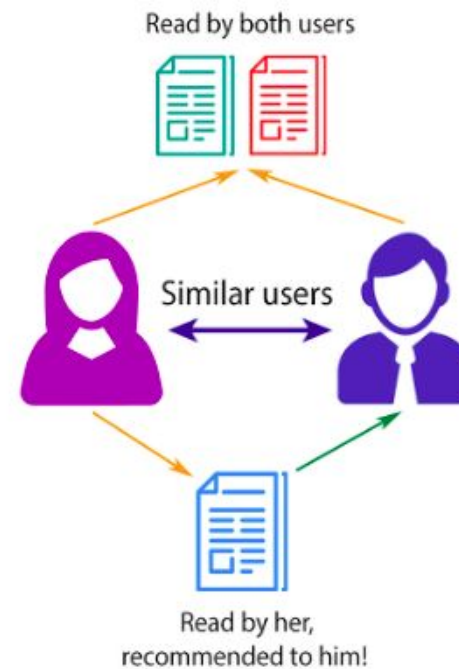
하이브리드 필터링

- 앞에서 다루었던 콘텐츠 기반 필터링과 협업 필터링을 다시
생각해보자.

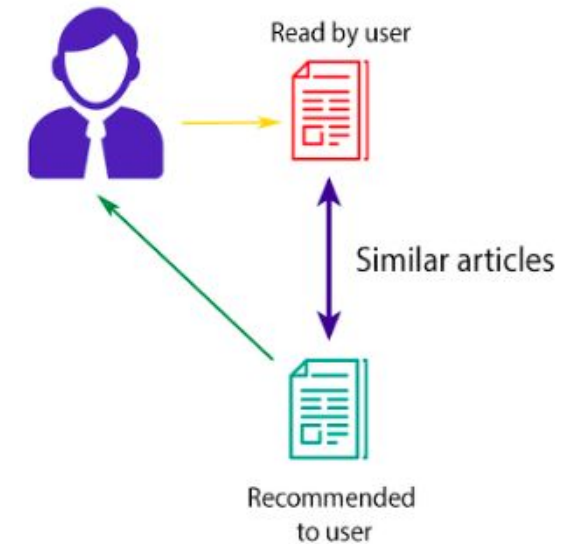
만약 리뷰데이터가 아직 없는 신규 제품이라면?



COLLABORATIVE FILTERING



CONTENT-BASED FILTERING



교육

Recommendation System

추천 시스템
종류

하이브리드 필터링

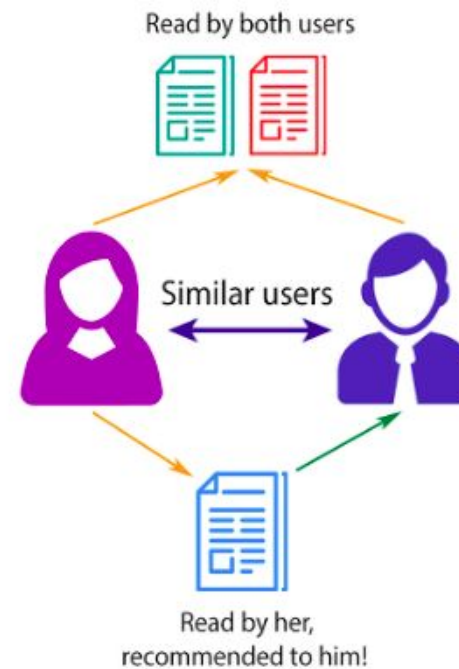
- 앞에서 다루었던 콘텐츠 기반 필터링과 협업 필터링을 다시
생각해보자.

만약 리뷰데이터가 아직 없는 신규 제품이라면?

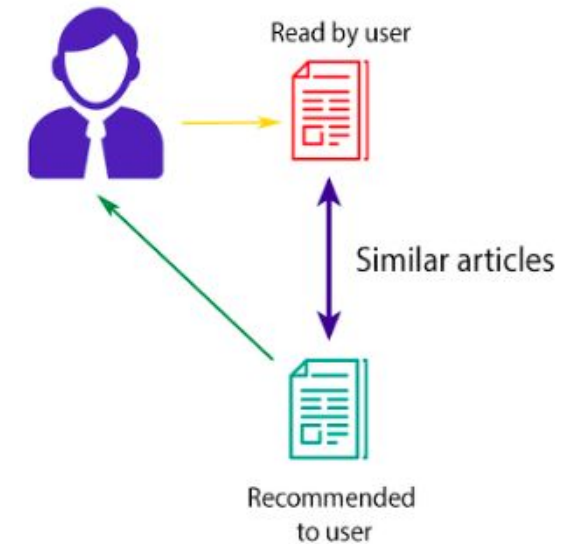
협업 필터링은 사용할 수 없다. **(콜드 스타트)**
하지만, 콘텐츠 기반 필터링은 사용할 수 있다.



COLLABORATIVE FILTERING



CONTENT-BASED FILTERING



교육

Recommendation System

추천 시스템
종류

하이브리드 필터링

- 앞에서 다루었던 콘텐츠 기반 필터링과 협업 필터링을 다시
생각해보자.

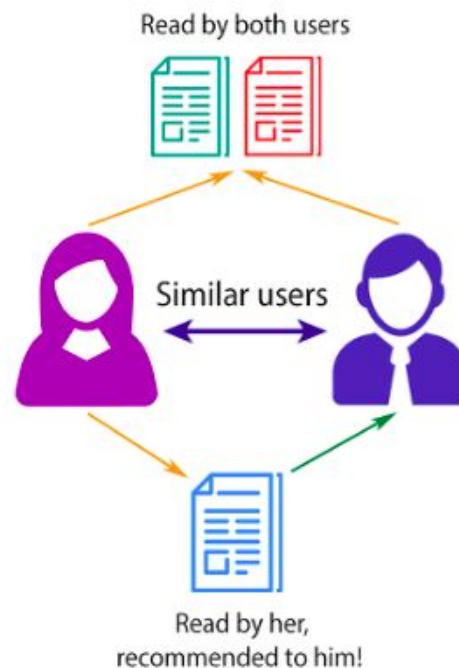
만약 리뷰데이터가 아직 없는 신규 제품이라면?

협업 필터링은 사용할 수 없다. **(콜드 스타트)**
하지만, 콘텐츠 기반 필터링은 사용할 수 있다.



그렇기 때문에 이 둘을 혼합하여,
신규 제품에 대해서는 콘텐츠 기반 필터링을 적용,
오래된 제품에 대해서는 사용자의 선호 데이터를 활용하는
협업 필터링을 사용하는 방식으로, 추천을 다채롭게 할 수 있다

COLLABORATIVE FILTERING



CONTENT-BASED FILTERING

