



aws INNOVATE

AI/ML EDITION

24 February 2022

기계학습 전문지식 없이 맞춤형 추천기능 개발하기

Amazon Personalize

김지선

테크니컬 트레이너

AWS



세부 주제

- 개인화 추천
- Amazon Personalize란?
- 구성 방법 및 주요기능
- Demo
- 활용예시와 도입 시 고려사항

개인화 추천

고객과의 모든 접점에 개인특성이 반영됩니다.

플레이리스트 순위

섹션 순서

카테고리

Shows



Movies

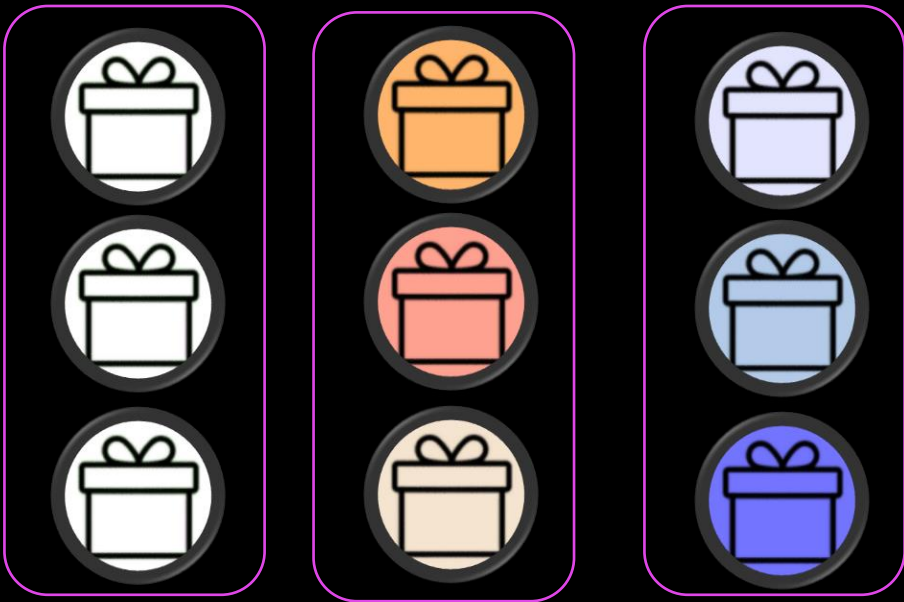


By genre



추천 알고리즘

- 콘텐츠 기반 필터링 (Content-based)
 - Item을 feature로 표현 후, item끼리 비슷한지 비교



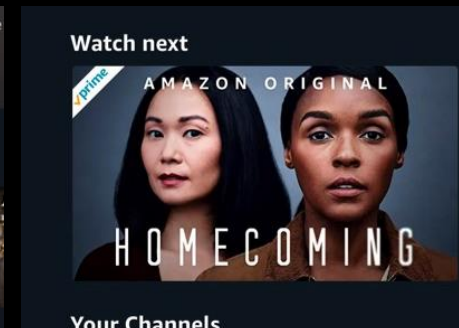
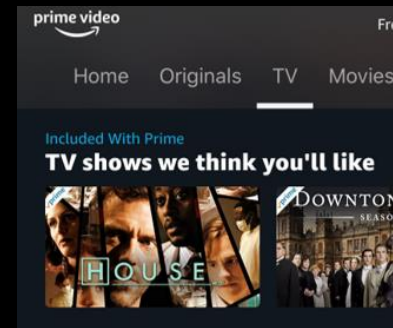
- 협업 필터링 (Collaborative)
 - User-item의 상호 작용(interaction) 정보를 통해 user가 좋아할 만한 item을 추천



개인화 & 추천의 트렌드

초개인화 Hyper Personalization

- 맞춤형 웹사이트 (추천 결과 기반 대시보드, A/B 테스트)
- AI기술 기반 개인화 (온라인 광고, 취향별 고객군 분석)
- 검색어 기반 실시간 추천 (소셜미디어)
- 개인화 분석을 돕는 증강현실기술

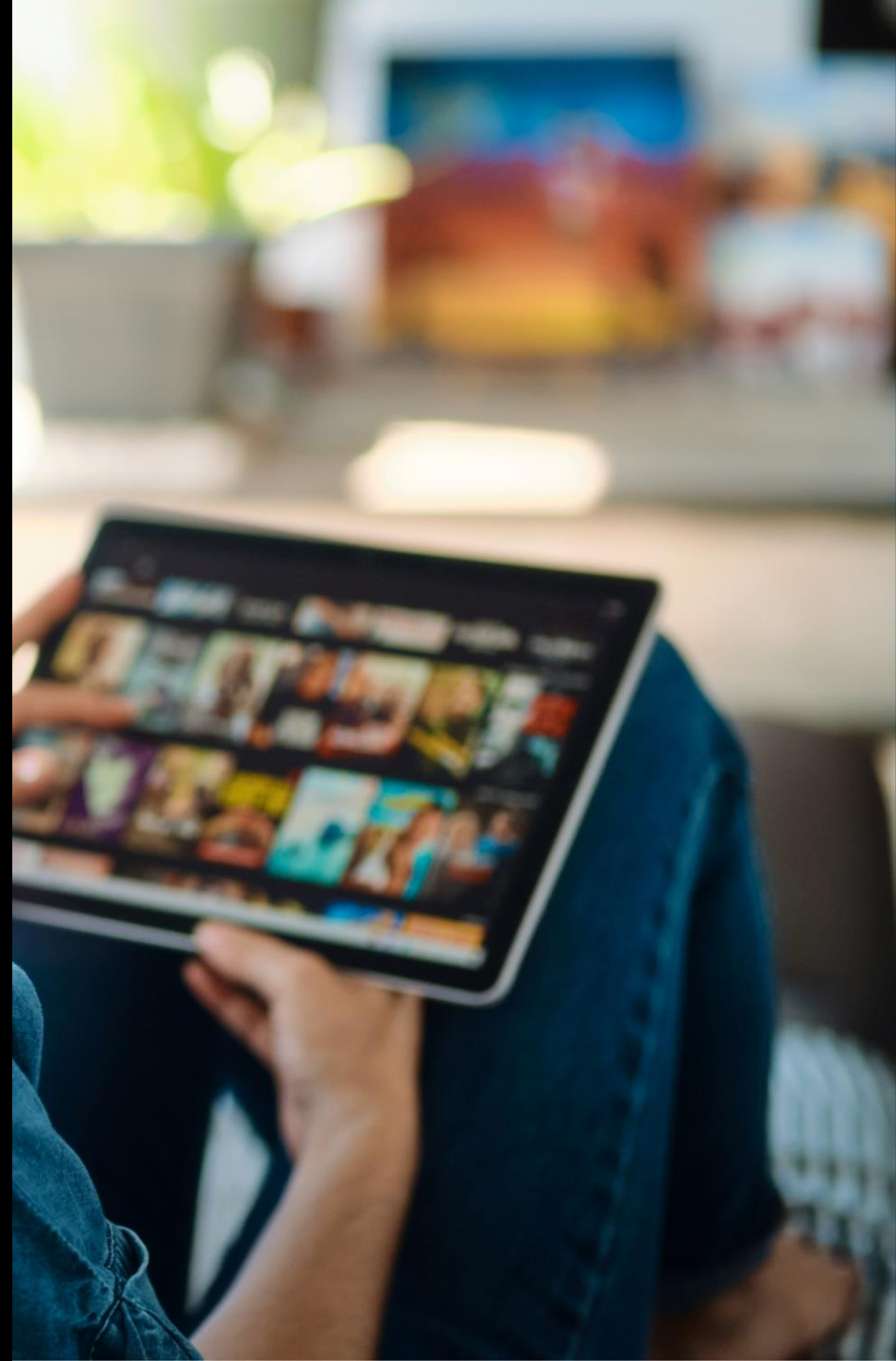


개인화 추천을 필수적으로 도입해야 하나요?

63%

소비자가 개인화를 표준서비스로
인식하고있습니다

*Source: <https://www.business2community.com/marketing/30-amazing-personalization-statistics-02289044>



개인화 추천의 과제



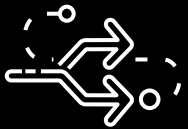
신규사용자/상품에 대한 데이터의 부족(Cold Start)



전문 인력 고용의 어려움



장기간의 개발



추천시스템 전반의 소유 및 관리에 대한 높은 비용

Amazon Personalize

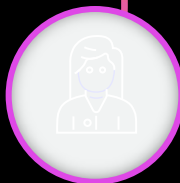
Amazon Personalize

amazon.com에서 사용하는 머신러닝 기술을 바탕으로 한 실시간 개인화 및 추천 서비스

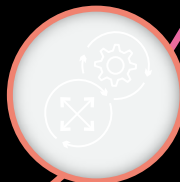
ML을 활용한 고객 경험 개선



Amazon Personalize가 클릭 몇 번으로 사용자 지정 ML 모델을 구축, 교육 및 배포하는 모든 작업을 수행하도록 하여 대규모 ML 솔루션을 쉽게 구현할 수 있습니다.



Api 기반의 추천 결과 생성 및 활용을 통해 기존의 웹, 앱, 이메일 마케팅 시스템에 쉽게 통합합니다.



모든 데이터는 비공개로 안전하게 암호화됩니다.
데이터는 고객 간에 또는 Amazon.com과 공유되지 않습니다.

Amazon Personalize

구성 방법 & 주요 기능

전체 워크 플로우

1

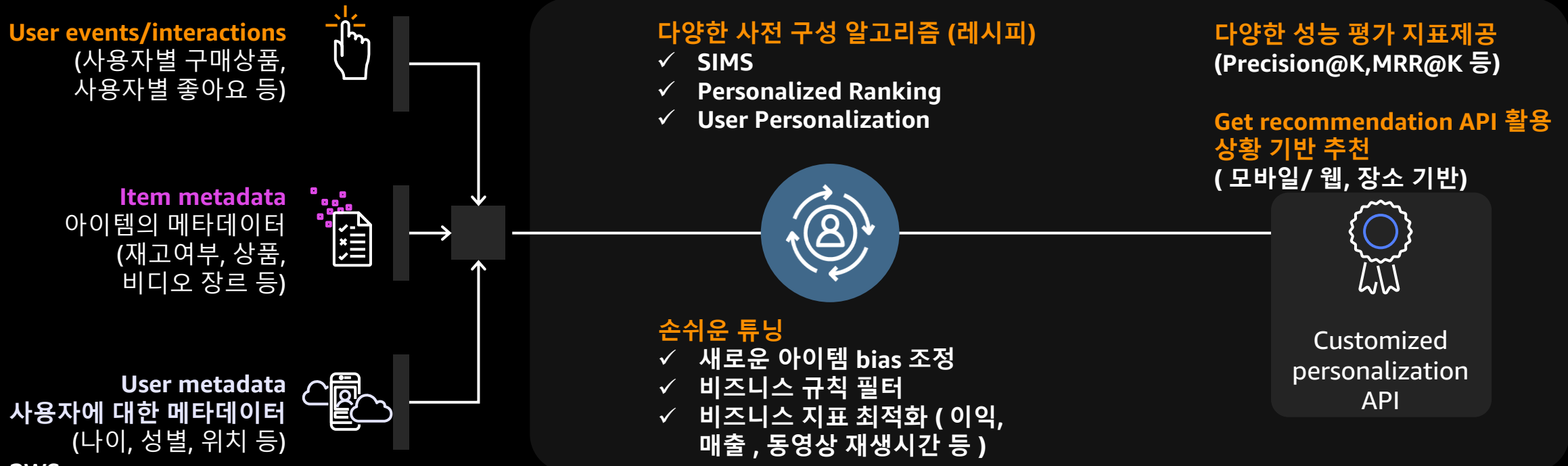
데이터 입력

2

솔루션 생성 및 튜닝

3

추천 결과 제공
모델 배포(캠페인)



1. 입력 데이터셋 준비

입력 데이터 스키마에 맞춘 데이터셋을 S3에 저장합니다.

○ Interactions

- ✓ User와 item 간의 인터랙션 정보 제공
- ✓ 예: 사용자별 구매상품, 사용자별 좋아요 등
- ✓ 모든 알고리즘에서 필수로 필요

[입력 예시]

user_id	item_id	timestamp
298	474	884182806

- ✓ Csv 포맷(첫번째라인은 컬럼헤더 정보)
- ✓ 최소규모의 데이터셋 준비 필요

○ Item

- ✓ 아이টে에 대한 메타데이터 제공
- ✓ 예: 재고여부, 상품, 비디오 장르 등

Item_id	Item_title
1	안경

○ User

- ✓ 사용자에게 대한 메타데이터 제공
- ✓ 예: 나이, 성별, 위치 등

user_id	username	age	gender
17	user17	18	F

* 보라색으로 표현된 부분은 필수적으로 필요한 스키마입니다.

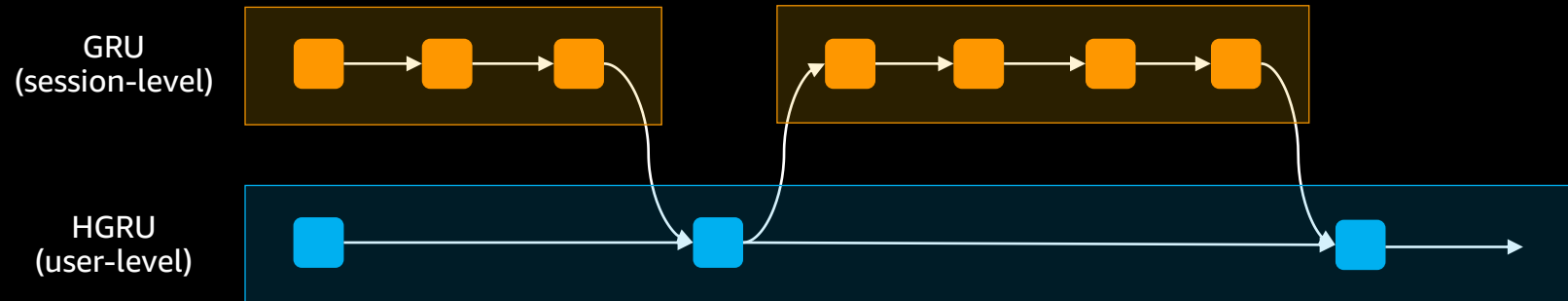
2. 솔루션 생성 - 알고리즘 (레시피)

레시피 유형	레시피
USER_PERSONALIZATION	User-Personalization Popularity-Count HRNN recipe (legacy) HRNN-Metadata recipe (legacy) HRNN-Coldstart recipe (legacy)
PERSONALIZED_RANKING	Personalized-Ranking
RELATED_ITEMS	Similar-Items SIMS
USER_SEGMENTATION	Item-Affinity Item-Attribute-Affinity

2. 솔루션 생성 - 알고리즘 (레시피)

아이템 추천 레시피

- User Personalization
 - 모든 시나리오에 적합하여 가장 추천되는 알고리즘
 - HRNN 알고리즘



Source: Session-based Recommendations with Recurrent Neural Networks (<https://arxiv.org/pdf/1511.06939.pdf>)

2. 솔루션 생성 - 알고리즘 (레시피)

아이템 추천 레시피

- Popularity count
 - 모든 사용자 행동 데이터를 기반으로 가장 인기 있는 항목을 추천
- Similar Items
 - 상호작용 데이터를 통해 유사 아이템 추천
 - 상품 탐색과 상세페이지 구성 시 활용
- Personalized ranking
 - 소비자의 예측 관심도에 따라 순서 순위로 추천 생성

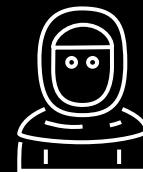
2. 솔루션 생성 – User Segmentation 레시피

2021
신규기능

지능형 고객 세분화 레시피

영화, 제품 등과 같은 특정 항목에 관심이 있는
사용자를 식별하여 그룹화합니다.

마케팅 채널을 통한 보다 효과적인 캠페인
새로운 상품 카테고리, 장르, 채널 등에 대한 사용자 확보
마케팅 지출에 대한 투자 수익 향상



Action movie fans

2. 솔루션 생성 – User Segmentation 레시피

지능형 고객 세분화 레시피

- ✓ Item data 필수입력데이터
- ✓ JSON포맷

[입력 예시]

- Item-Affinity

```
{"itemId":"1"}  
{"itemId":"2"}  
{"itemId":"3"}  
{"itemId":"4"}  
{"itemId":"5"}
```

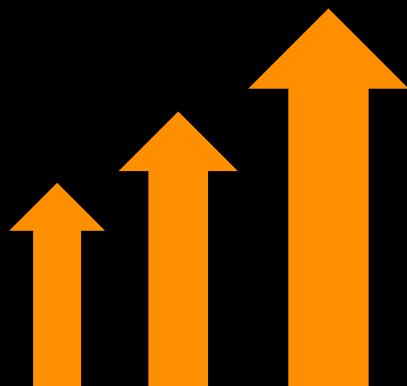
- Item-Attribute-Affinity

```
{"itemAttributes":  
  "ITEMS.genres = \"Comedy\"  
  AND ITEMS.genres =  
    \"Action\""}  
}
```

3. 솔루션 평가- 성능평가지표(metrics)

다양한 성능 평가 지표를 통해 최적의 솔루션을 선별할 수 있습니다.

- 다양한 성능 평가 지표를 제공
- 가장 관련도가 높은 아이템에 높은 스코어를 부여하는 알고리즘(recipe) 선정 시 유용



지표	설명
Precision@K	$\frac{\text{실제 테스트 데이터와 일치하는 추천 아이템 개수}}{\text{추천 아이템 전체 개수}}$
MRR@K (Mean Reciprocal Rank)	모든 쿼리에 대한 상위 25개 권장 사항에서 첫 번째 관련 권장 사항의 역 순위 평균입니다. 이 지표는 가장 높은 순위의 단일 권장 사항에 관심이 있는 경우에 적합합니다.
NDCG@K (Normalized Discounted Cumulative Gain)	이상적인 추천 결과로부터 구할 수 있는 최대 가능 스코어를 기준으로 정규화 한, 관련 아이템 순위에 역 로그 가중치를 반영하여 계산함
Coverage	학습 데이터의 전체 고유한 아이템 수 중에서 실제 쿼리에서 추천된 아이템 수

4. 솔루션 배포 - 캠페인 생성

- 솔루션을 지정 및 배포하여 캠페인을 생성
- 업데이트 방식
 - ✓ 자동 : 솔루션이 업데이트 될 때마다 가장 최신 버전의 솔루션으로 자동배포
 - ✓ 수동 : updatecampaign을 call 해서 캠페인을 수동으로 업데이트
- status가 active로 변경 된 후 캠페인 작업 사용 가능

4. 솔루션 배포 - 테스트 캠페인 결과



- 임의의 사용자 아이디를 입력하면 해당 사용자에게 추천하는 상위 25개의 아이템을 확인할 수 있습니다.

Test campaign results

User ID [Info](#)
This is the user ID of the user you want to see campaign results for. This user ID needs to be obtained from your user-interactions or user dataset.

Filter name- optional
Choose an existing filter to apply to your user segments or create a new filter.

None ▼

  [Create new filter](#)

Context - optional
Provide context as key/value pairs.

Key	Value	
<input type="text" value="Enter a metadata attribute"/>	<input type="text" value="Enter value"/>	<input type="button" value="Remove"/>

5. 캠페인 활용

캠페인 생성 후 추천 결과 활용 (API 호출)

레시피 유형	API	API 요구 사항
USER_PERSONALIZATION	GetRecommendations	userId: 필수 itemId: 사용되지 않음 inputList: 해당 사항 없음
PERSONALIZED_RANKING	GetPersonalizedRanking	userId: 필수 itemId: 해당 사항 없음 inputList: itemId 목록
RELATED_ITEMS	GetRecommendations	userId: 사용되지 않음 itemId: 필수 inputList: 해당 사항 없음
USER_SEGMENTATION	CreateBatchSegmentJob	Batch segmentation job 생성

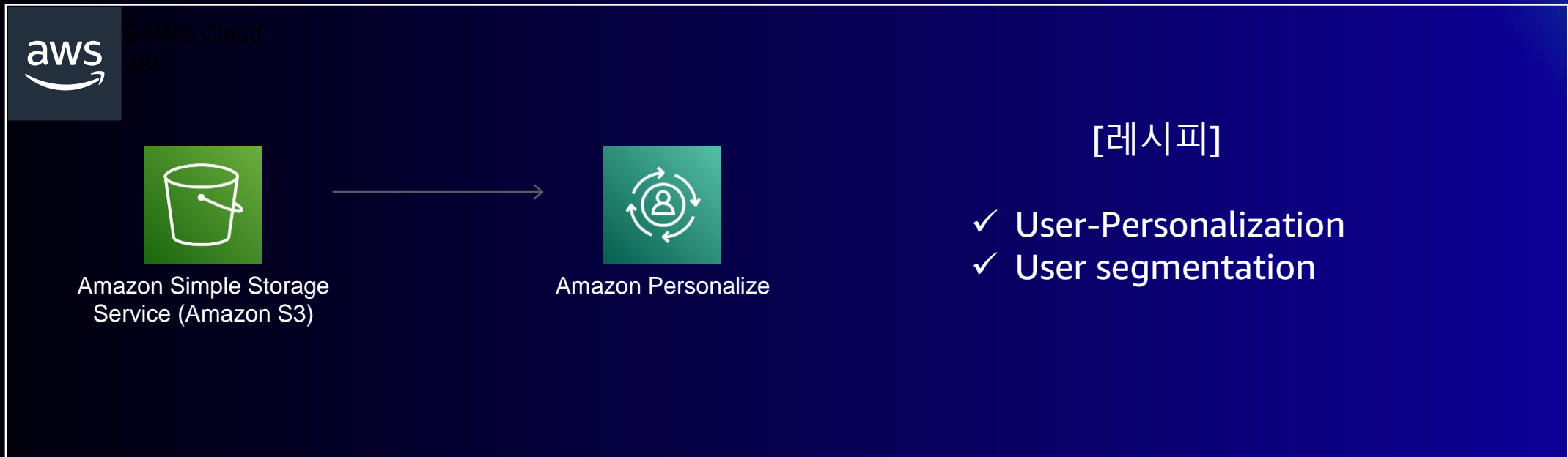
Demo



Demo

원본 데이터 :

<https://github.com/aws-samples/amazon-personalize-with-aws-glue-sample-dataset>





main

1 branch

0 tags

Go to file

Code



zoishp Update README.md

8769915 on Feb 12, 2021 3 commits



CODE_OF_CONDUCT.md

Initial commit

11 months ago



CONTRIBUTING.md

Initial commit

11 months ago



LICENSE

Initial commit

11 months ago



README.md

Update README.md

11 months ago



item.csv

adding files

11 months ago



user-item-interaction.json

adding files

11 months ago



README.md

Setting up Amazon Personalize with AWS Glue Sample Dataset

This dataset is to be used with the AWS blog "Setting up Amazon Personalize with AWS Glue".

item.csv is a list of ecommerce items to be used with Amazon Personalize

user-item-interaction.json is a list of users purchase patterns in relation to items.csv. It is to be used with AWS Glue to prepare it for Amazon Personalize.

Security

See [CONTRIBUTING](#) for more information.

About

No description, website, or topics provided.



Readme



View license



Code of conduct



2 stars



2 watching



2 forks

Releases

No releases published

Packages

No packages published

Contributors 2



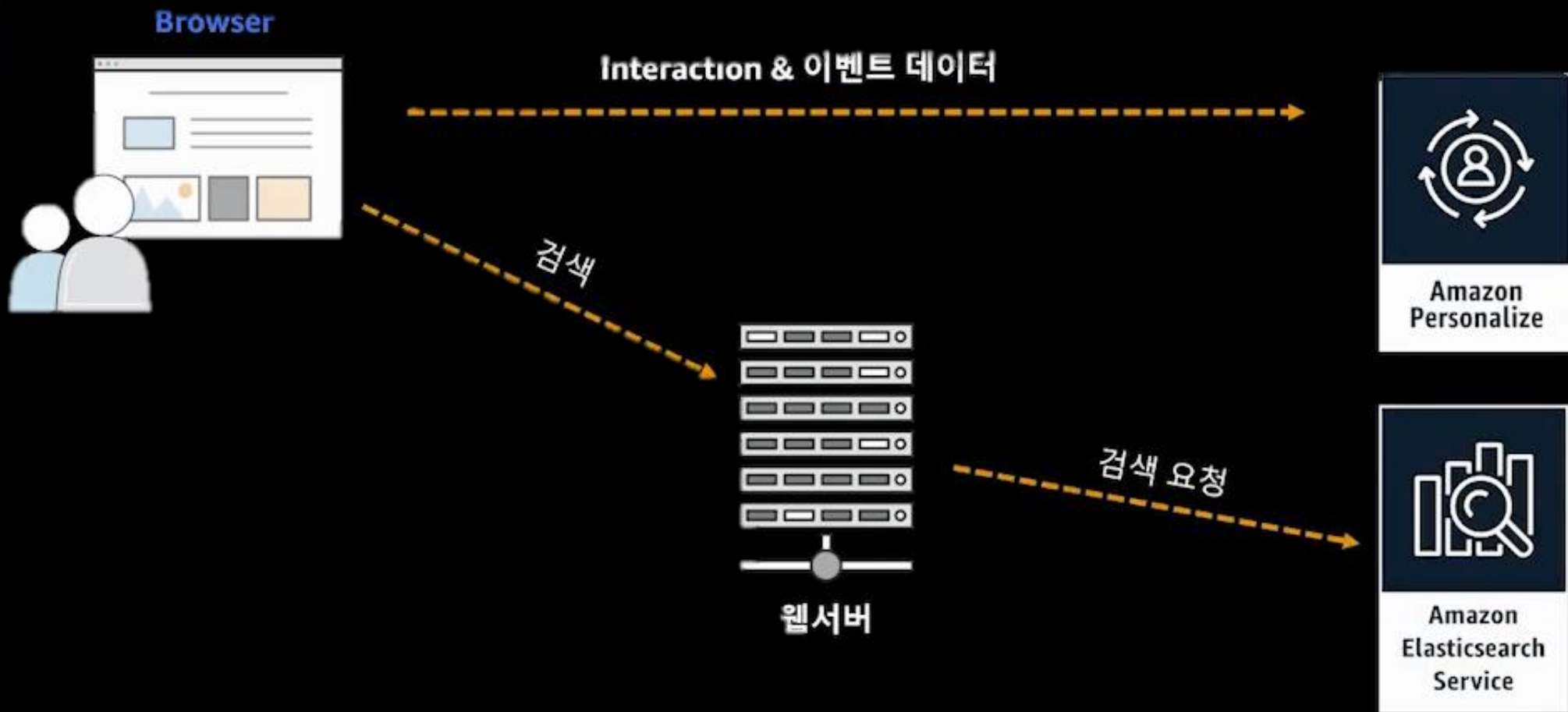
zoishp Zoish Pithawala



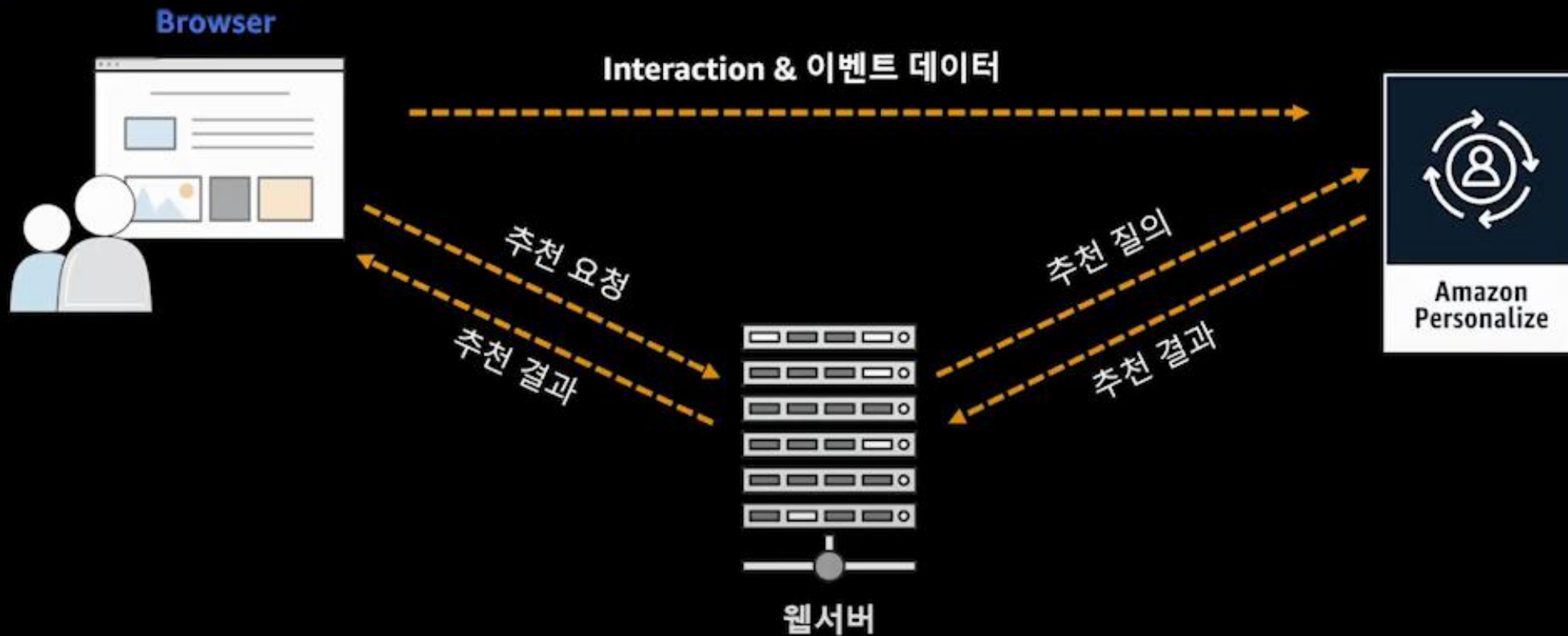
amazon-auto Amazon GitHub Autom...

활용 예시와 도입 시 고려사항

활용 예시- 개인화 검색



활용 예시 - 실시간 추천



도입 시 고려사항

Good Fit

- ✓ 사용자에게 상품을 추천
- ✓ 유사 상품 카탈로그
- ✓ 개인화 마케팅
- ✓ 개인화 대시보드 구성

Bad Fit

- ✓ 적은 데이터 사이즈
- ✓ 대부분의 사용자가 익명의 사용자
- ✓ 단순한 규칙기반의 추천

“전세계 모든 데이터 과학자와 개발자가
손쉽게 활용할 수 있는
편리한 인공지능 & 머신러닝 환경을 제공”

The vision of AWS for Machine Learning & Artificial Intelligence

AI & ML 리소스 허브

AWS가 제공하는 AI 및 ML에 관한 다양한 자료들을 통해 더욱 심층적으로 학습해보세요!

- 기계 학습 여정 가이드
- 기계 학습의 7가지 주요 사용 사례
- 데이터, 분석 및 기계 학습을 위한 전략 플레이북
- 올바른 클라우드 서비스 및 인프라를 통한 기계 학습 혁신 가속화 전략 가이드
- 기계 학습에 적합한 컴퓨팅 인프라 선택 가이드
- 컨택트 센터의 서비스 개선 및 비용 절감 방법
- + 외의 다양한 동영상 학습 자료 및 기술 학습 자료!



<https://bit.ly/3yUk0Kx>

리소스 허브 방문하기

AWS Innovate - AI/ML 특집에 참석해주셔서 대단히 감사합니다.

저희가 준비한 강연, 어떻게 보셨나요?
더 나은 세미나를 위하여 **설문을 꼭 작성해 주세요!**



aws-korea-marketing@amazon.com



twitter.com/AWSKorea



facebook.com/amazonwebservices.ko



youtube.com/user/AWSKorea



linkedin.com/company/amazon-web-services



twitch.tv/aws

Thank you!