

# Personalize Pipeline APIs(ver 21.01.26)

## API

- Private API Gateway 를 통해 API 가 제공되므로 로컬에서는 BastionHost 등을 통해 접근해야한다.
- 모든 API 는 Content-type:application/json 을 기본으로 사용한다.
- 아래의 표에서 볼드처리된 파라미터는 필수 파라미터이다.

### Step-functions integrated

step function 으로 통합된 API 는 비동기방식으로 호출 된다.  
종료된 결과는 notify.py 를 통해 이메일 및 슬랙으로 알람이 온다.

#### POST /personalize/cleanup

데이터셋 그룹과 관련된 모든 리소스(데이터셋, 캠페인, 솔루션, 스키마)를 지운다.

Param	Type	Example	Description
<b>name</b>	string	hmn-movielens	데이터셋 그룹 이름

#### POST /personalize/[user-personalization|sims|ranking]

각각 User Personalization, SIMS, Personalized-Ranking 레시피를 이용하여 캠페인을 생성한다.

schema\_arn 은 [GET|POST] /personalize/schema 를 통해 얻을 수 있다.

Param	Type	Example	Description
<b>name</b>	string	movielens	생성할 데이터셋 그룹 이름
<b>bucket</b>	string	s3://dongkyl-test/xyz.csv	interactions 데이터의 s3 path
<b>schema_arn</b>	string	arn:aws:personalize:.... /test-event-schema	스키마 arn
<b>perform_hpo</b>	string		1 파라미터에 값이 존재하면 HPO 를 이용하여 솔루션을 생성한다
<b>solution_config</b>	object	<pre>"solution_config": {   "algorithmHyperParameters": {     "min_cointeraction_count": "0",     "popularity_discount_factor":       "0.43490596145165816"   } }</pre>	하이퍼파라미터 등 옵션

#### POST /personalize/batch-inference

배치로 추론을 한다. HRNN 류 레시피는 input json 값은 아래와 같은 형태로 입력한다. (SIMS 는 itemId)

```
{  
  "userId": "38"  
}  
{  
  "userId": "663"  
}  
{  
  "userId": "3384"  
}
```

Param	Type	Example	Description
-------	------	---------	-------------

<b>name</b>	string	movielens	생성할 배치 인퍼런스 잡 이름
<b>solution_version_arn</b>	string	arn:aws:personalize:ap-northeast-2:776556808198:solution/metadata-hpo/5c699484	솔루션 버전 ARN
<b>input_path</b>	string	s3://dongkyl.personalize.batch/input.json	stacked json 형태 인풋값
<b>output_path</b>	string	s3://dongkyl.personalize.batch/output/	결과가 저장될 위치

## POST /personalize/hrnn-metadata

HRNN-Metadata 레시피를 이용하여 캠페인을 생성한다.

item\_bucket 또는 user\_bucket 둘 중 하나는 반드시 입력되어야 한다.

각 schema\_arn 은 [GET|POST] /personalize/schema 를 통해 얻을 수 있다.

Param	Type	Example	Description
<b>name</b>	string	movielens	생성할 데이터셋 그룹 이름
<b>bucket</b>	string	s3://dongkyl-test/xyz.csv	interactions 데이터의 s3 path
<b>schema_arn</b>	string	arn:aws:personalize:..../inter-schema	interactions 스키마 arn
<b>item_bucket</b>	string	s3://dongkyl-test/xyz.csv	items 메타데이터의 s3 path
<b>item_schema_arn</b>	string	arn:aws:personalize:..../item-schema	items 메타데이터의 스키마 arn
<b>user_bucket</b>	string	s3://dongkyl-test/xyz.csv	users 데이터의 s3 path
<b>user_schema_arn</b>	string	arn:aws:personalize:..../user-schema	users 메타데이터의 스키마 arn
<b>perform_hpo</b>	string		1 파라미터에 값이 존재하면 HPO 를 이용하여 솔루션을 생성한다
<b>solution_config</b>	object	"solution_config": { "algorithmHyperParameters": { "min_cointeraction_count": "0", "popularity_discount_factor": "0.43490596145165816" } }	하이퍼파라미터 등 옵션

## POST /personalize/metadata-dataset

데이터셋 임포트는 기존의 데이터를 replace 하는 것이다. 증분식 업데이트는 불가능하다.

기존에 존재하는 hrnn(또는 metadata) 데이터셋 그룹에 메타데이터 데이터셋을 추가하거나 임포트하고 싶을때 사용한다.

item\_bucket & item\_schema\_arn 또는 user\_bucket & user\_schema\_arn 둘 중 하나는 반드시 입력되어야 한다.

1. 지정한 메타데이터 데이터셋들을 추가하고 (이미 존재하면 데이터셋 임포트 잠만 추가함)
2. 솔루션과 캠페인을 추가한다

콘솔상에서 데이터셋과 스키마 는 강결합 되어 있다. (즉 데이터셋 생성시 사용한 스키마만 사용해야 함.)

Param	Type	Example	Description
<b>name</b>	string	movielens	기존에 존재하는 데이터셋 그룹 이름
<b>item_bucket</b>	string	s3://dongkyl-test/xyz.csv	items 메타데이터의 s3 path
<b>item_schema_arn</b>	string	arn:aws:personalize:..../item-schema	items 스키마 arn
<b>user_bucket</b>	string	s3://dongkyl-test/xyz.csv	users 메타데이터의 s3 path
<b>user_schema_arn</b>	string	arn:aws:personalize:..../user-schema	users 스키마 arn
<b>perform_hpo</b>	string		1 파라미터에 값이 존재하면 HPO 를 이용하여 솔루션을 생성한다
<b>solution_config</b>	object	"solution_config": { "algorithmHyperParameters": { "min_cointeraction_count": "0", "popularity_discount_factor": "0.43490596145165816" } }	하이퍼파라미터 등 옵션

		} }	
--	--	--------	--

## POST /personalize/interactions-dataset

데이터셋 임포트는 기존의 데이터를 replace 하는 것이다. 증분식 업데이트는 불가능하다.

기존에 존재하는 데이터셋 그룹에 Interactions 데이터셋을 추가하거나 임포트하고 싶을때 사용한다.

1. Interactions 데이터셋을 추가하고 (이미 존재하면 데이터셋 임포트 잡만 추가함)
2. 솔루션과 캠페인을 추가한다

콘솔상에서 데이터셋과 스키마 는 강결합 되어 있다. (즉 데이터셋 생성시 사용한 스키마만 사용해야 함.)

Param	Type	Example	Description
name	string	movielens	기존에 존재하는 데이터셋 그룹 이
bucket	string	s3://dongkyl-test/xyz.csv	interactions 데이터의 s3 path
schema_arn	string	arn:aws:personalize::.../schema	interactions 스키마 arn
recipe_arn	string	arn:aws:personalize:::recipe/aws-sims	솔루션 생성에 사용할 recipe arn
perform_hpo	string		1 파라미터에 값이 존재하면 HPO 를 이용하여 솔루션을 생성한다
solution_config	object	"solution_config": { "algorithmHyperParameters": { "min_cointeraction_count": "0", "popularity_discount_factor": "0.43490596145165816" } }	하이퍼파라미터 등 옵션

## POST /personalize/train-recipe

지정된 레시피를 이용하여 현재 데이터셋을 학습시킨다.

Param	Type	Example	Description
name	string	movielens	기존에 존재하는 데이터셋 그룹 이
recipe_arn	string	arn:aws:personalize:::recipe/aws-sims	솔루션 생성에 사용할 recipe arn
perform_hpo			
solution_config	object	"solution_config": { "algorithmHyperParameters": { "min_cointeraction_count": "0", "popularity_discount_factor": "0.43490596145165816" } }	하이퍼파라미터 등 옵션

## Lambda Integrated

람다 통합된 API 는 엔드포인트와 비슷한 이름을 가진 람다함수 하나와 매핑된다.

## GET /personalize/tracking

PutEvents 에서 사용할 트래킹 아이디를 가져온다.

이벤트 트래커가 존재해야 한다.(데이터셋 목록에서 직접 생성 할 수 있고, 파이프라인을 통해서 생성했다면 자동으로 생성된다.)

Param	Type	Example	Description
name	string	hrnn-movielens	데이터셋 그룹의 이름

## GET /personalize/metrics

생성된 캠페인에 대해 학습시 평가한 성능을 보여준다.

Param	Type	Example	Description
name	string	hrnn-movielens	캠페인 이름

## GET /personalize/campaigns

현재 생성된 모든 캠페인의 arn 목록을 보여준다.

## GET /personalize/recommend/sims

sims 레시피로 학습된 캠페인을 쿼리하여 추천 아이템 목록을 가져온다.

Param	Type	Example	Description
campaign_arn	string	arn:aws:personalize:...	캠페인의 arn
item_id	string	242	아이템 아이디
num_results	integer	25	반환될 아이템의 개수

## GET /personalize/recommend/hrnn

hrnn 또는 hrnn-metadata 레시피로 학습된 캠페인을 쿼리하여 추천 아이템 목록을 가져온다.

Param	Type	Example	Description
campaign_arn	string	arn:aws:personalize:...	캠페인의 arn
user_id	string	dongkyl	사용자 아이디
num_results	integer	25	반환될 아이템의 개수

## GET /personalize/recommend/ranking

personalized-ranking 레시피로 학습된 캠페인을 쿼리하여 재정렬된 아이템 목록을 가져온다.

Param	Type	Example	Description
campaign_arn	string	arn:aws:personalize:...	캠페인의 arn
user_id	string	dongkyl	사용자 아이디
item_list	string	22,34,56	csv 아이템 목록

## POST /personalize/schema

지정된 JSON 내용으로 스키마를 생성한다.

Param	Type	Example	Description
name	string	my-schema	스키마 이름
		{ "type": "record", "name": "Users", "namespace": }	

schema	string	<pre> namespace : "com.amazonaws.personalize.schema", "fields": [ { "name": "USER_ID", "type": "string" }, { "name": "AGE", "type": "int" }, { "name": "GENDER", "type": "string", "category": true } ], "version": "1.0" } </pre>	JSON 문자열 형 태의 스키마 (json stringified string)
--------	--------	--	--

## GET /personalize/schema

현재 생성된 모든 schema 목록을 가져온다.

## Kinesis Integrated

PutEvents API 는 기본이 1000 rps 를 처리하게 되어 있으며, 한 요청에 10개 까지만 이벤트를 보낼 수 있다. 따라서 동기방식으로 처리하면 최대 1000 rps, 비동기로 버퍼를 이용하면 최대 10000rps 요청이 가능하다. 최소한으로 요청보내기 위해 키네시스를 이용하여 스트림에 쌓고 비동기로 PutEvents 를 호출한다.

## POST /personalize/put-events

퍼스널라이즈의 PutEvents API 를 통해 실시간으로 이벤트를 입력해준다.  
event\_type, user\_id, item\_id 는 학습에 사용된 데이터와 같은 포맷이어야 한다.

Param	Type	Example	Description
tracking_id	string	movielens	GET events 를 통해 얻은 tracking_id
session_id	string	대략 uuid	임의로 생성한 세션의 아이디
event_type	string	Click	이벤트 타입
user_id	string	dongkyl	사용자의 아이디
item_id	string	242	아이템 아이디