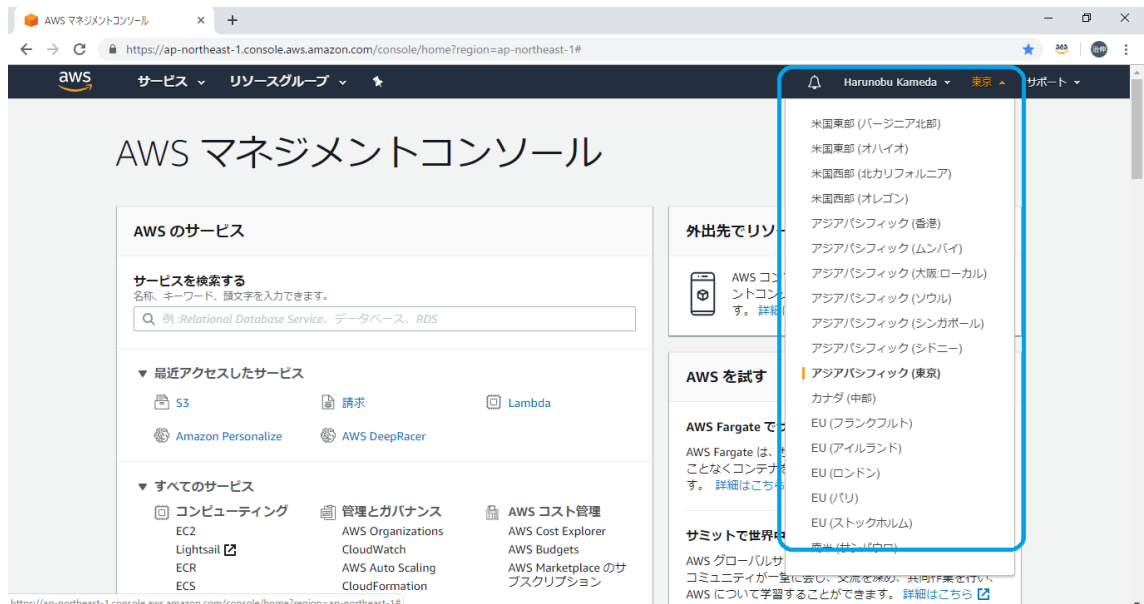


Amazon Personalize ハンズオン (No.Codes)

2019/08/15

アマゾン ウェブ サービス ジャパン エバンジェリスト 亀田

1. ハンズオンを行うリージョンを画面右上から選択。リージョンは講師の指示に従ってください。



2. 映画の視聴履歴ファイルをダウンロード

<https://bit.ly/2XEEQxk>

ユーザー610 人による 9700 本の映画視聴履歴とその評価が 10 万件以上ファイルに含まれています。

USER_ID : ユーザーの識別 ID (商用環境では別途ユーザーマスタと連結

してください)

ITEM_ID : 映画の ID (商用環境では別途アイテムマスタと連結してください)

TIMESTAMP : 映画批評を入力された時間です。(このハンズオンでは用いませんが、Amazon Personalize の実行では必須なカラムです)

Amazon Personalize はこれらのデータをもとに、批評の相関図を作成しレコメンドを行います。従来これらのデータ分析は、Amazon EMR などを用いたヒューリスティック分析が必要でしたが、Amazon Personalize ではこれらが自動化されることが特徴です。

例) User (1) が 映画 A、B、C を視聴した場合、User(2)が B を視聴した場合、A と C がレコメンドされる。レコメンド以外にも、データの相関関係を見出す用途で様々な用途が今後想定されます。また、本シナリオでは使用しませんが、CSV ファイルには Rating (視聴済映画に対するユーザーからの評価) も準備されており、ユーザー評価に基づいたさらに細かいレコメンドが Amazon Personalize では可能です。

3. Amazon S3 バケットに先ほど DL した ratings.csv をアップロードします。S3 バケットは既存のものを作成してもよいですし、新規で作成しても問題ありません。S3 バケットは以下のバケットポリシーを設定してくだ

さい。すでにバケットポリシーが設定されている既存 S3 バケットでの作業
はお勧めしていません。

```
{  
  
  "Version": "2012-10-17",  
  
  "Id": "PersonalizeS3BucketAccessPolicy",  
  
  "Statement": [  
  
    {  
  
      "Sid": "PersonalizeS3BucketAccessPolicy",  
  
      "Effect": "Allow",  
  
      "Principal": {  
  
        "Service": "personalize.amazonaws.com"  
  
      },  
  
      "Action": [  
  
        "s3:GetObject",  
  
        "s3:ListBucket"  
  
      ],  
  
      "Resource": [  
  
        "arn:aws:s3:::bucket-name",
```

```
"arn:aws:s3:::bucket-name/*"

    ]

}

]

}
```

bucket-name はご自身で作業するバケットの名前を入れてください。

S3 バケットを新規に作成する場合は、全てデフォルトで作成してください。

【バケットを作成する】のボタンを押します。



バケット名は全 AWS ユーザーの全リージョンで一意である必要がありますので、ある程度長い名称にします。リージョンは、講師から指示されたリージョンを指定してください。



バケットが作成されたらバケット名をダブルクリックして詳細画面へ遷移し、【アクセス権限】タブを押します。



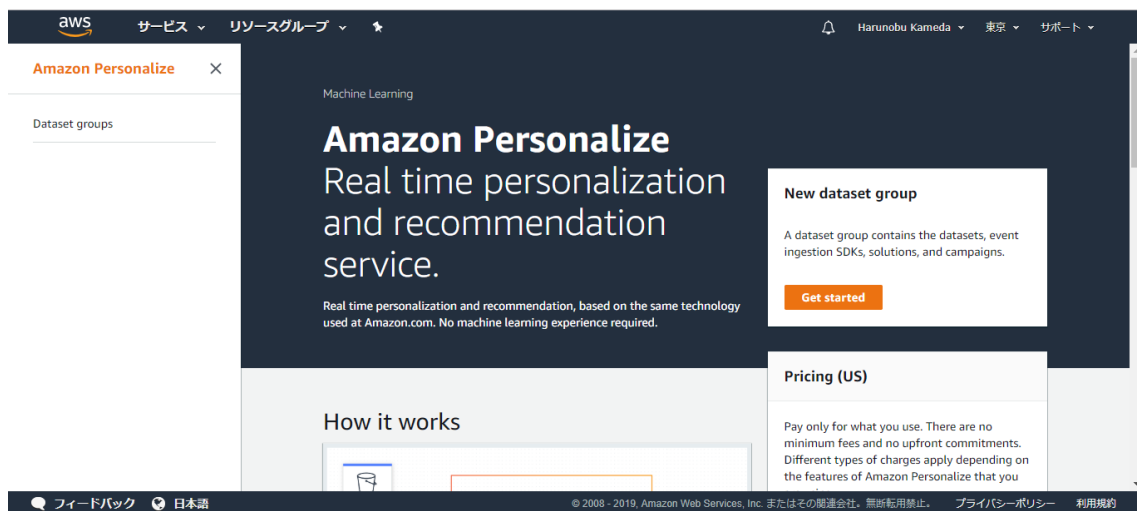
【バケットポリシー】をボタンを押し、上記のポリシー内容をコピーし、【保存】を押します。



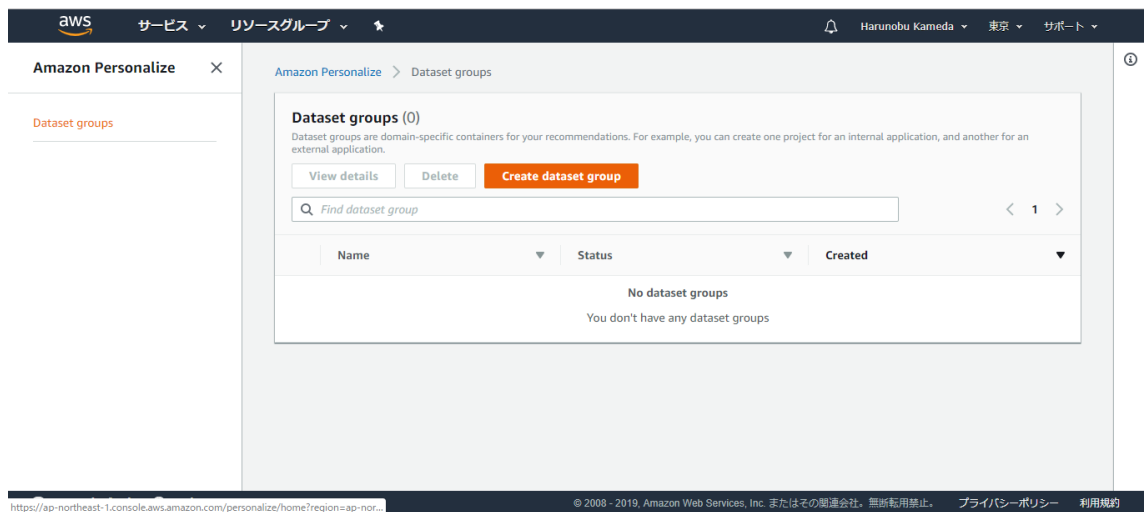
4. Amazon Personalize のマネージメントコンソール

(<https://console.aws.amazon.com/personalize/>) にアクセスします。

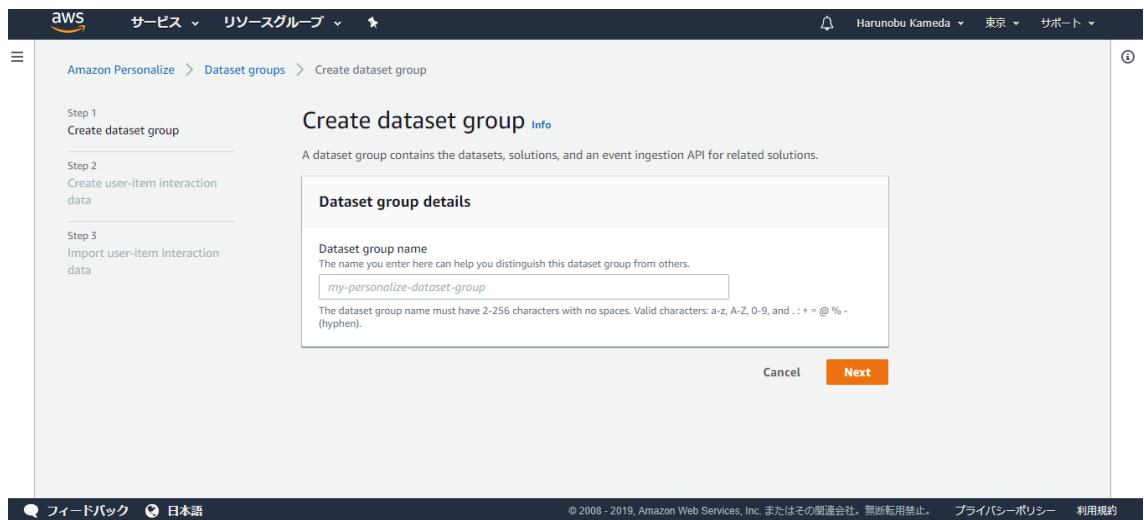
先ほど指定したリージョンになっていることを改めて確認します。



5. 画面左の [Dataset groups]を選択し、[Cerate dataset group]を押してください。



6. 作成する Dataset Group に任意の名前を付けてください。



7. 作成する Dataset に名前を付けます

aws サービス リソースグループ

Harunobu Kameda 東京 サポート

Step 1 Create dataset group

Step 2 Create user-item interaction data

Step 3 Import user-item interaction data

Create user-item interaction data

Now that you have created your dataset group, Amazon Personalize requires user-item interaction data to create a solution. The first step in creating your user-item interaction dataset is to provide Amazon Personalize with the schema of the dataset. The schema you provide allows Amazon Personalize to understand and import your dataset.

Dataset details

Dataset name
The name you enter here can help you distinguish this dataset import job from others.

20190701personalize

The dataset name must have 2-256 characters with no spaces. Valid characters: a-z, A-Z, 0-9, and . : + = @ % - (hyphen).

Schema details

Schema selection

☐ Use existing schema
Choose an existing schema that matches your dataset.

☒ Create new schema
Create a new schema to match your dataset

フィードバック 日本語 © 2008 - 2019, Amazon Web Services, Inc. またはその関連会社。無断転用禁止。 プライバシーポリシー 利用規約

8. [Create new schema]を選択し、Schema に任意の名前を設定します。

aws サービス リソースグループ

Harunobu Kameda 東京 サポート

Schema selection

☐ Use existing schema
Choose an existing schema that matches your dataset.

☒ Create new schema
Create a new schema to match your dataset

New schema name
The name you enter here appears in the Schema dashboard. It can help you distinguish this schema from others.

20190701personalize

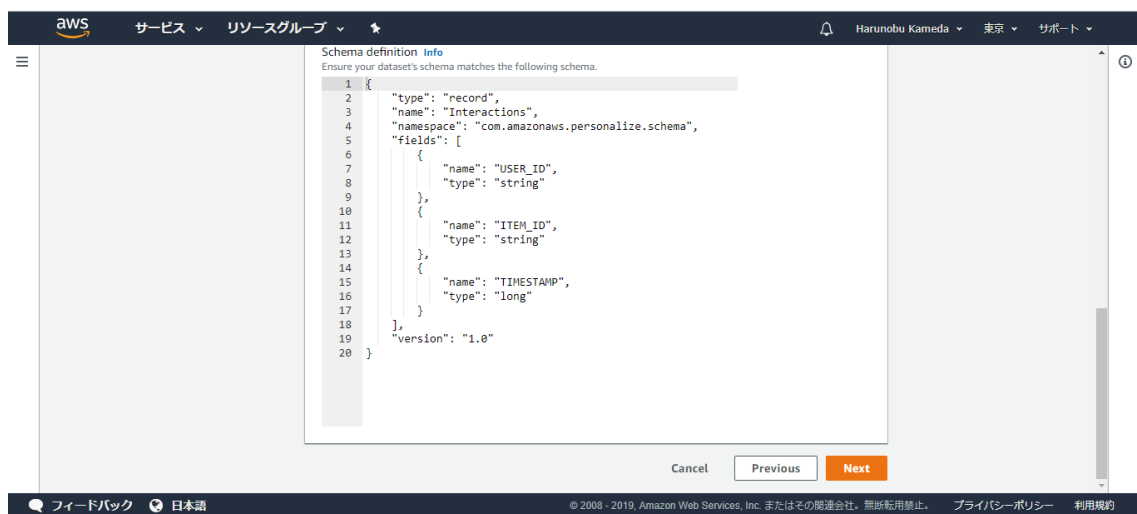
The schema name must have 2-256 characters with no spaces. Valid characters: a-z, A-Z, 0-9, and . : + = @ % - (hyphen).

Schema definition
Ensure your dataset's schema matches the following schema.

```
1 {
2   "type": "record",
3   "name": "Interactions",
4   "namespace": "com.amazonaws.personalize.schema",
5   "fields": [
6     {
7       "name": "USER_ID",
8       "type": "string"
9     },
10    {
11      "name": "ITEM_ID",
12      "type": "string"
13    },
14    {
15      "name": "TIMESTAMP",
16      "type": "string"
17    }
18  ]
19 }
```

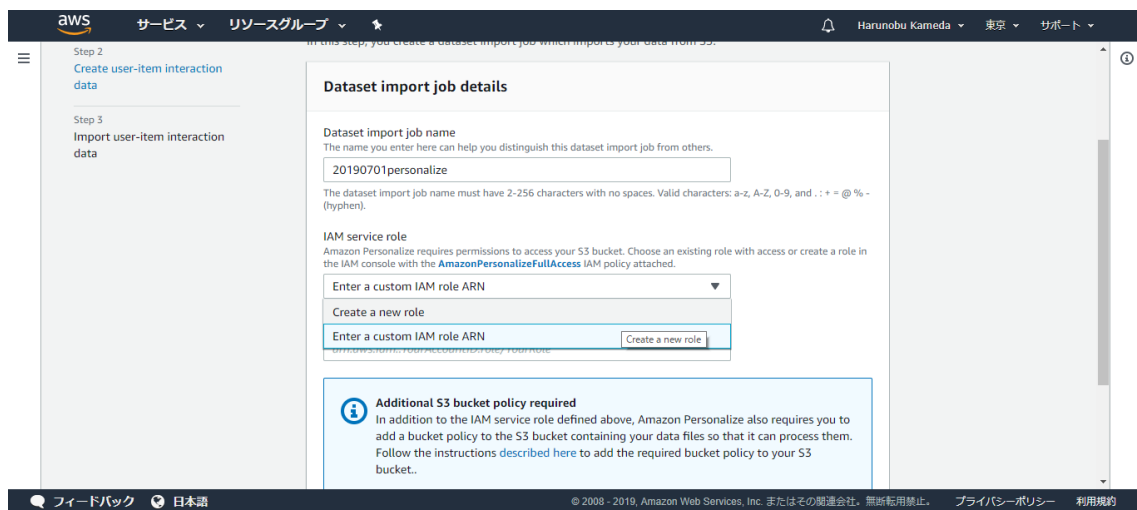
フィードバック 日本語 © 2008 - 2019, Amazon Web Services, Inc. またはその関連会社。無断転用禁止。 プライバシーポリシー 利用規約

9. [Next]を押します



10. [user-item interaction data] (先ほどの ratings.csv)をインポートする job の起動を行います。[Dataset import job name]に名前を入力します。

[IAM service role]で[Create a new role]を選択します。



11. 先ほど作成した S3 バケットの名前を[Specify S3 buckets]に入力して[Create role]を押します。

Create an IAM role ✕

Passing an IAM role gives Amazon Personalize permission to perform actions in other AWS services on your behalf. Creating a role here will grant permissions described by the [AmazonPersonalizeFullAccess](#) IAM policy to the role you create.

The IAM role you create will grant access to:

- ☒ S3 buckets you specify - optional
 - ☒ Specific S3 buckets

20190701personalize

Comma delimited. ARNs, "*" and "/" are not supported.
 - ☐ Any S3 bucket
 Allow users that have access to your notebook instance access to any bucket and its contents in your account.
 - ☐ None
- ☒ Any S3 bucket with "personalize" in the name
- ☒ Any S3 object with "Personalize" in the name
- ☒ S3 bucket with a Bucket Policy allowing access to Amazon Personalize [See S3 bucket policy](#)

Cancel Create role

1 2. [Success! You created an IAM role.]が表示されていることを確認します。

[Data location]に `s3://<bucketname>/ratings.csv` を入力します。

Amazon Personalize ✕

Dataset groups

- ▼ 20190701personalize
 - Dashboard
 - Datasets
 - Event trackers [Info](#)
 - Solutions and recipes [Info](#)
 - Campaigns [Info](#)

IAM service role
Amazon Personalize requires permissions to access your S3 bucket. Choose an existing role with access or create a role in the IAM console with the [AmazonPersonalizeFullAccess](#) IAM policy attached.

AmazonPersonalize-ExecutionRole-1561960763200

☒ **Success! You created an IAM role.**
AmazonPersonalize-ExecutionRole-1561960763200

Additional S3 bucket policy required
In addition to the IAM service role defined above, Amazon Personalize also requires you to add a bucket policy to the S3 bucket containing your data files so that it can process them. Follow the instructions [described here](#) to add the required bucket policy to your S3 bucket.

Data location [Info](#)
Choose the S3 location of your data.

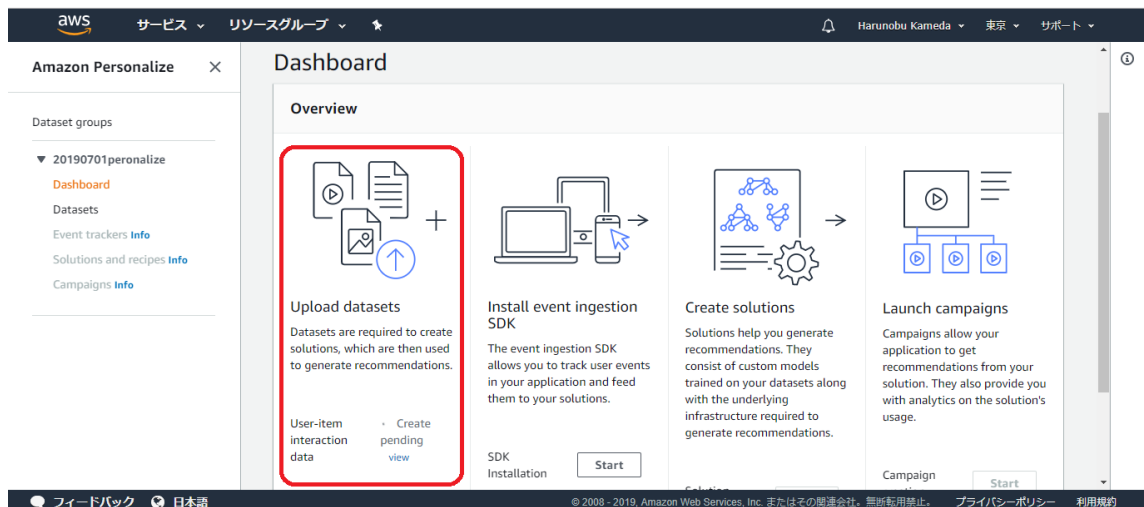
s3://20190701personalize/ratings.csv

Your file needs to be in a CSV format and reflect the schema.

Cancel Previous Finish

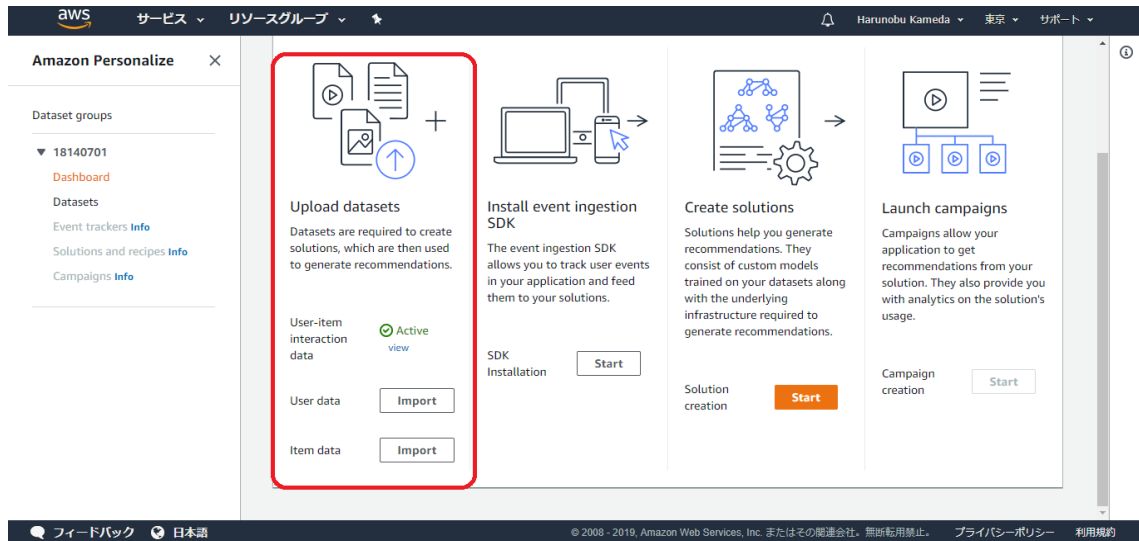
1 3. ステータスが[Create Pending] or [Create in Progress]になりますので少し待ちます。

Finish ボタンを押しても画面が遷移しない場合、ブラウザ上部の見えないところでエラーが出ていますので、右側のブラウザスクロールバーで一番上まで戻りエラーメッセージを確認してください。(ほとんどのケースにおいて、S3 のパス、もしくはファイル名のタイプミスです)

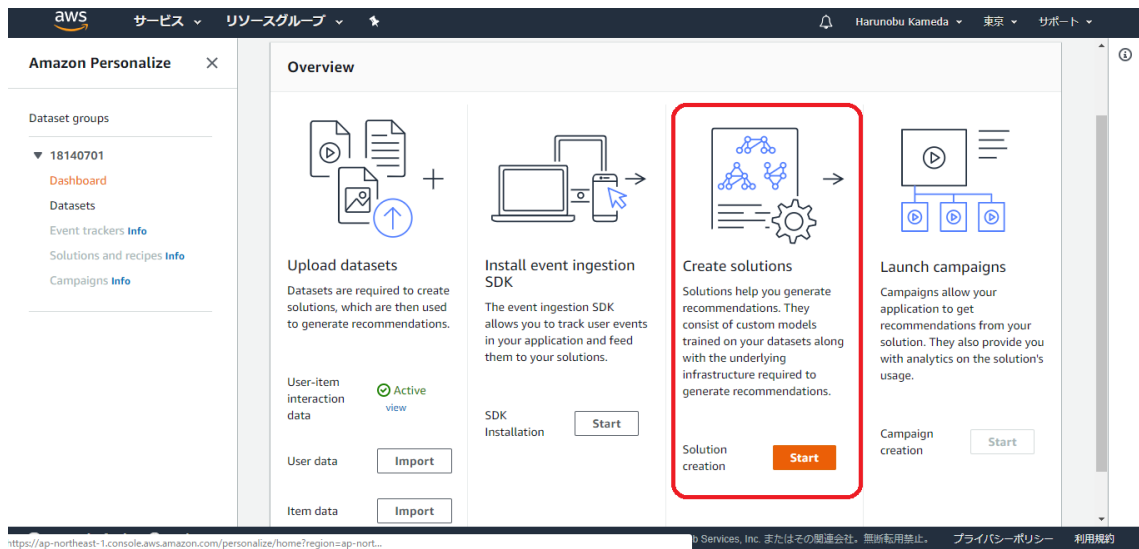


1 4. 取り込んだデータ(ratings.csv)の読み込みに少し時間がかかります。(20 分ぐらい)

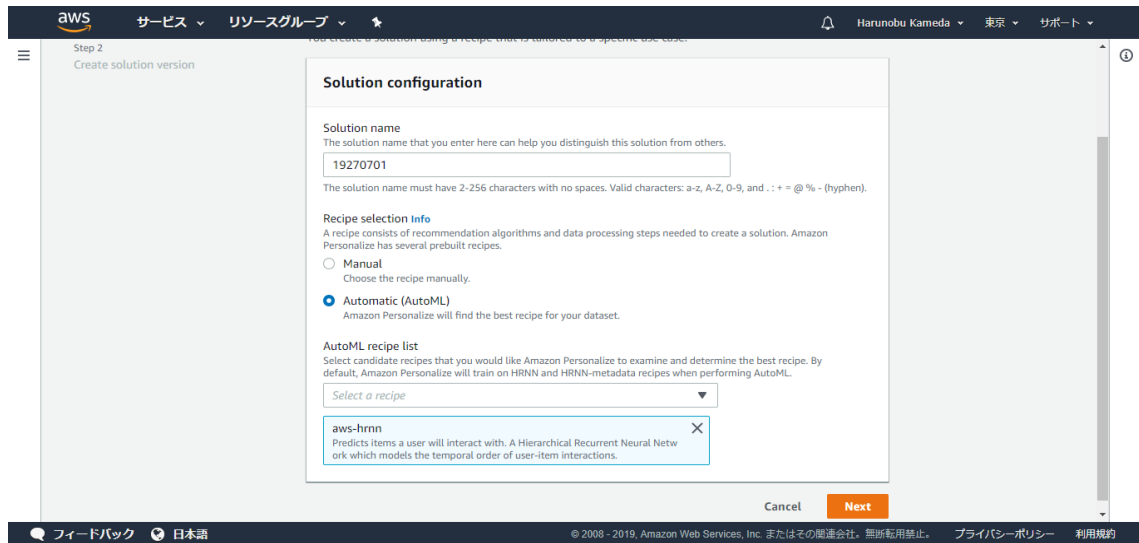
[Active]と表示されれば完了です。



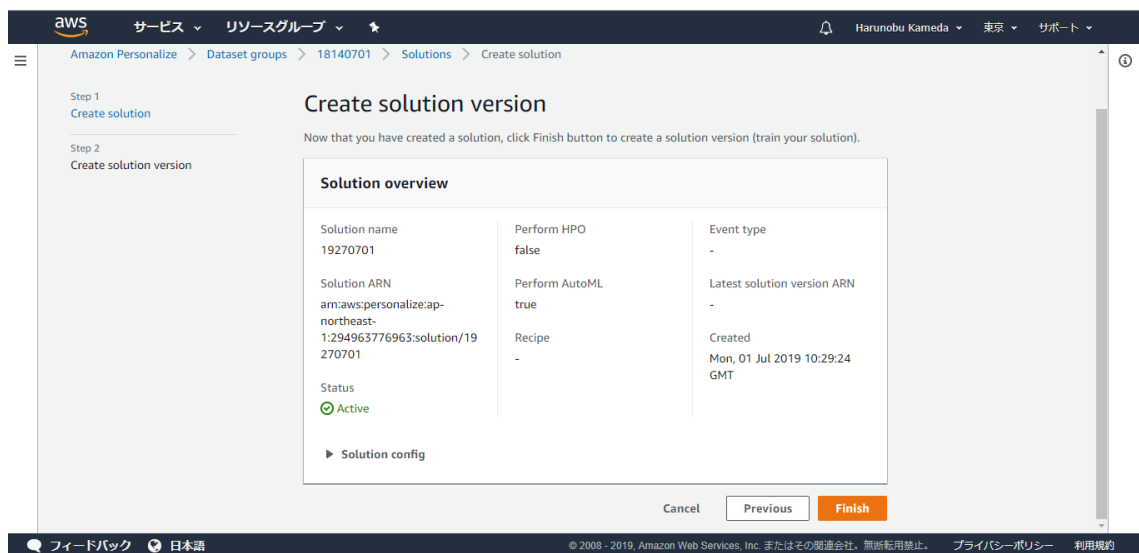
1 5. [Create Solutions]の[Start]を押します。読み込んだ datasets をもとに学習を開始します。



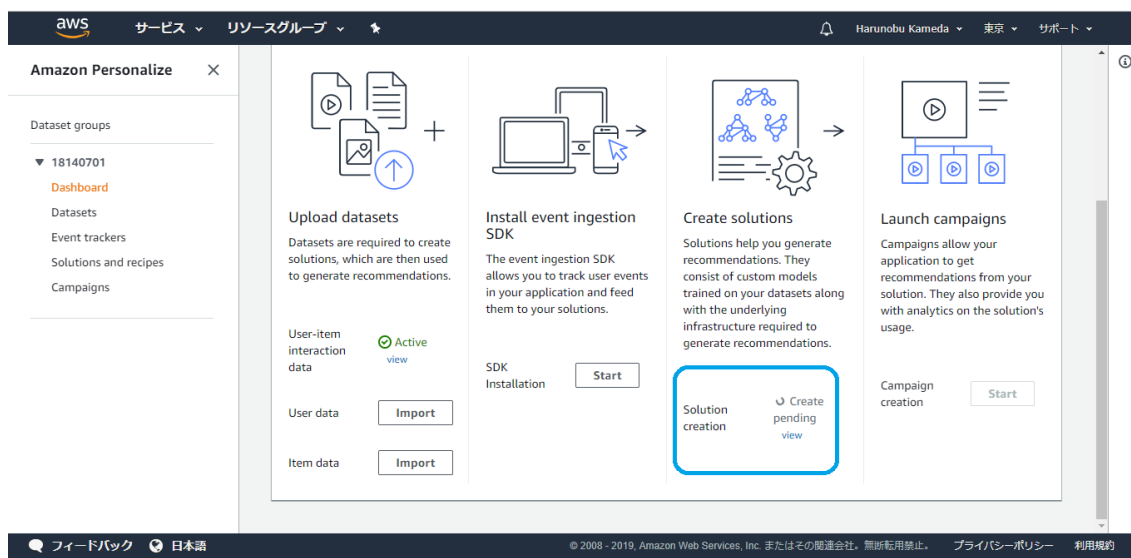
1 6. [Solution Name]を入力し、[Automatic (AutoML)]を選択します。
ドロップダウンリストから[aws-hrnn]を選択し以下の画面のようになれば、[Next]を押します。



1 7. 以下の画面で[Finish]を押すと学習が開始されます。



1 8. 学習は 50 分から 1 時間ぐらいかかります。



待ち時間の間に：

以下の AWS の AI サービス系を試してみましょう。

作業に戻ってきやすいようにブラウザ別のタブから管理者画面にアクセスします。

Amazon Rekognition：画像および動画の識別。

Amazon Polly：音声合成

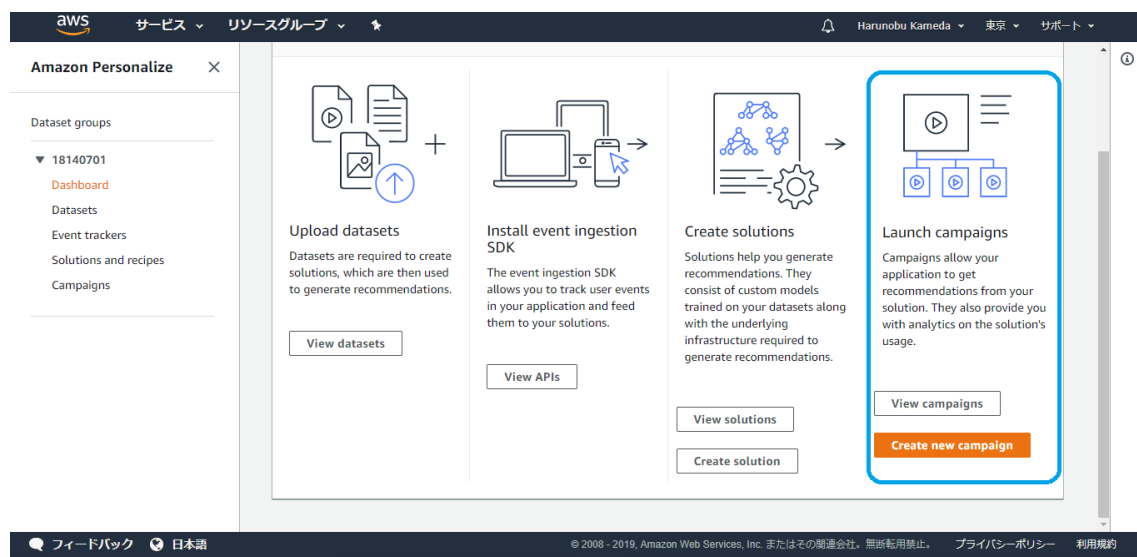
Amazon Translate：機械翻訳

3つのサービスとも、ハンズオンなしで管理者画面から操作することで基本的な機能を試すことができるようになっています。Rekognition はご自身の顔写真を jpeg ないしは png 形式でアップロードし、識別をしてみてください。

また、AWS CLI が使える方は巻末の手順を参考に AWS CLI の実行環境の準備

をしてみてください。

19. 学習が完了すると、[Create New Campaign]のボタンを押せるようになります。Campaign は一般的な機械学習用語でいうと、推論モデルのデプロイになります。



20. [Campaign Name]に任意の名前を入れます。[Solution]は先ほど作成した Solution の名前をドロップダウンから選択します。[Minimum provisioned transactions per second]はデフォルト状態のまま[Create Campaign]を押します。

aws サービス リソースグループ

Amazon Personalize > Dataset groups > 18140701 > Campaigns > Create new campaign

Create new campaign

Campaign details

Campaign name
The text you enter here appears in the Campaign dashboard and detail page. It can help you distinguish this campaign from others.

The campaign name must have 2-256 characters with no spaces. Valid characters: a-z, A-Z, 0-9, and . : + = @ % - (hyphen).

Solution
The selected solution is used to generate the recommendations provided in your campaign.

Minimum provisioned transactions per second [Info](#)
The minimum amount of throughput in transactions per second (TPS) that is provisioned for this campaign.

Enter a number from 1-500.

Cancel **Create campaign**

フィードバック 日本語 © 2008 - 2019, Amazon Web Services, Inc. またはその関連会社。無断転用禁止。 プライバシーポリシー 利用規約

エラーが表示された場合、ソリューションの詳細画面を開いてください。

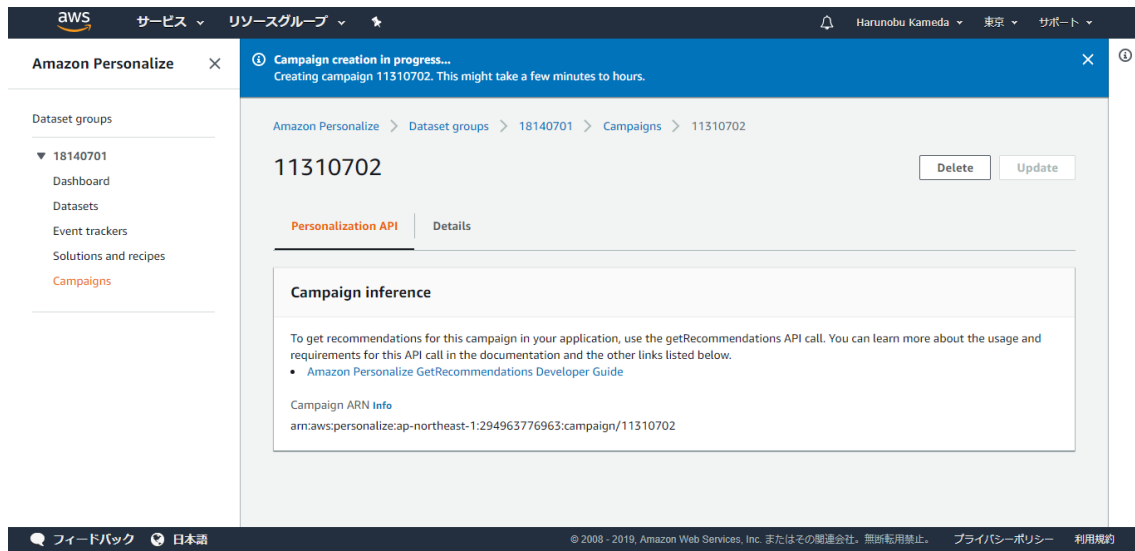
Solution versions (2)

Find solution version

Solution version ID	Solution version status	Created
8b48359f	Active	Thu, 15 Aug 2019 13:31:06 GMT
8e54b2a0	Active	Thu, 15 Aug 2019 13:31:03 GMT

ソリューションは自動で作成されますが、1 個が作成中の可能性があります。

2 1. 推論環境の構築が終わるまで待ちます。



2 2. 作成が完了すると[Test Campaign Result]が表示されますので、1 から 601 までの任意のユーザーID を入力し、[Get Recommendations]を押します。ratings.csv をもとに、入力されたユーザーの映画視聴履歴と相関のある別ユーザーの映画視聴履歴からレコメンドが出力されます。

映画は ID で出力されますので以下のデータをもとに補正します。

<https://grouplens.org/datasets/movielens/>

<http://files.grouplens.org/datasets/movielens/ml-latest-small.zip>

その他：時間に余裕があれば LightSail で AWS CLI 経由での操作に挑戦してみてください。

[GetRecommendations]を用いて以下のようにレコメンドを出力できます。

```
aws personalize-runtime get-recommendations ¥
```

```
--campaign-arn arn:aws:personalize:us-west-2:acct-
```

```
id:campaign/MovieRecommendationCampaign ¥
```

```
--user-id 123
```

以下のように JSON でレコメンドが出力されます。

```
{
```

```
  "itemList": [
```

```
    {
```

```
      "itemId": "14"
```

```
    },
```

```
    {
```

```
      "itemId": "15"
```

```
    },
```

```
    {
```

```
      "itemId": "275"
```

```
    },
```

```
    {
```

```
      "itemId": "283"
```

```
    },
```

```
{  
  "itemId": "273"  
},  
...  
]  
}
```

一番簡単な AWS CLI 環境の構築は Amazon Lightsail です。

1. Amazon Lightsail の管理画面で[インスタンスの作成]を押します



2. [OS のみ]を選択します



2. Amazon Linux が選択されていることを確認し、残りは全てデフォルト値で[インスタンスの作成]を押します。

3. インスタンスが作成中となりますので少し待ちます



4. 実行中になったら起動が完了ですので、コンソールボタンを押します。



5. [sudo pip install -U awscli]で CLI をアップグレードします。

6. [aws configure]でデフォルトリージョンを作業中のリージョンに指定します。(東京であれば[ap-northeast-1])

CLI の実行権限([AmazonPersonalizeFullAccess]と

[AmazonPersonalizeFullAccess]を合わせて設定します。)この際設定する

IAM ロールは[AmazonPersonalizeFullAccess]を保持している必要があります。

[Default output format]はデフォルトのまま[None]を設定します。

7. aws personalize-runtime get-recommendations --campaign-arn

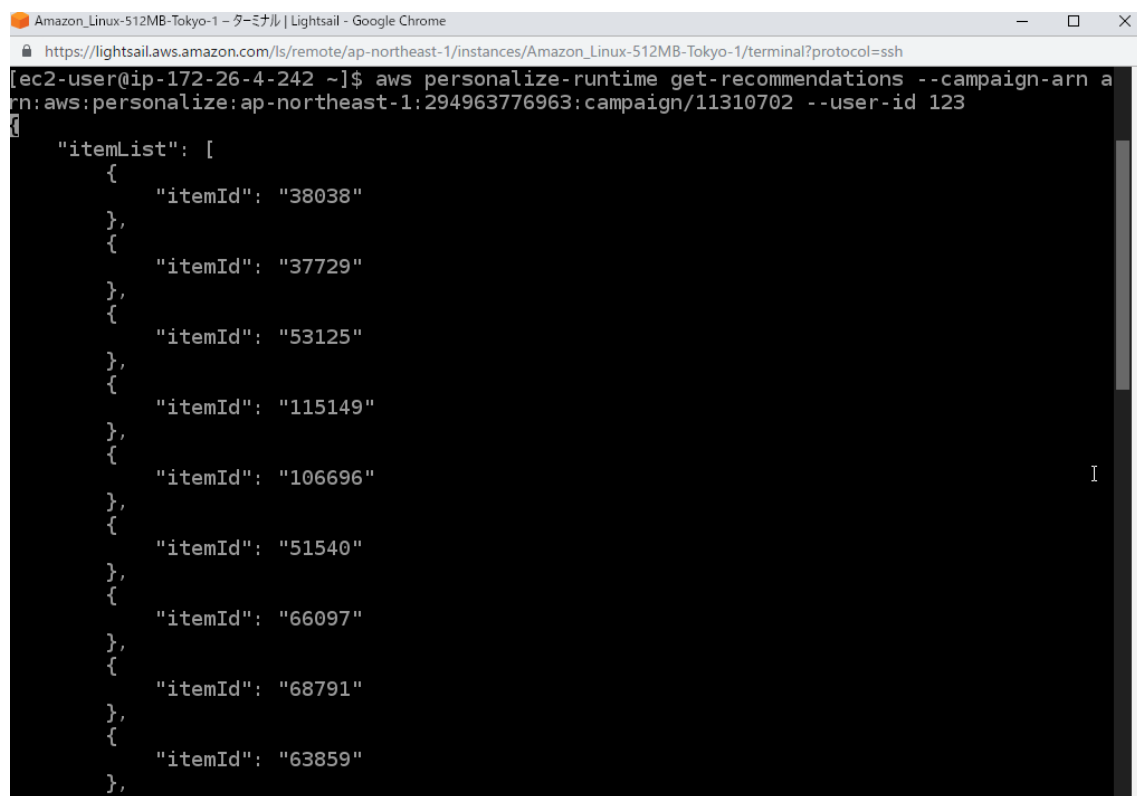
arn:<personalize campaign の arn> --user-id 123

のように入力し実行します。

以下のようなエラーがでた場合、IAM ロールの設定ミスが発生しています。

An error occurred (AccessDeniedException) when calling the
GetRecommendations operation: User:
arn:aws:sts::320976636470:assumed-
role/AmazonLightsailInstanceRole/i-0fa8803af5e4611ee is not
authorized to access this resource

8. 成功した場合以下が表示されます。



```
Amazon_Linux-512MB-Tokyo-1 - ターミナル | Lightsail - Google Chrome
https://lightsail.aws.amazon.com/ls/remote/ap-northeast-1/instances/Amazon_Linux-512MB-Tokyo-1/terminal?protocol=ssh
[ec2-user@ip-172-26-4-242 ~]$ aws personalize-runtime get-recommendations --campaign-arn arn:aws:personalize:ap-northeast-1:294963776963:campaign/11310702 --user-id 123
{
  "itemList": [
    {
      "itemId": "38038"
    },
    {
      "itemId": "37729"
    },
    {
      "itemId": "53125"
    },
    {
      "itemId": "115149"
    },
    {
      "itemId": "106696"
    },
    {
      "itemId": "51540"
    },
    {
      "itemId": "66097"
    },
    {
      "itemId": "68791"
    },
    {
      "itemId": "63859"
    }
  ]
}
```