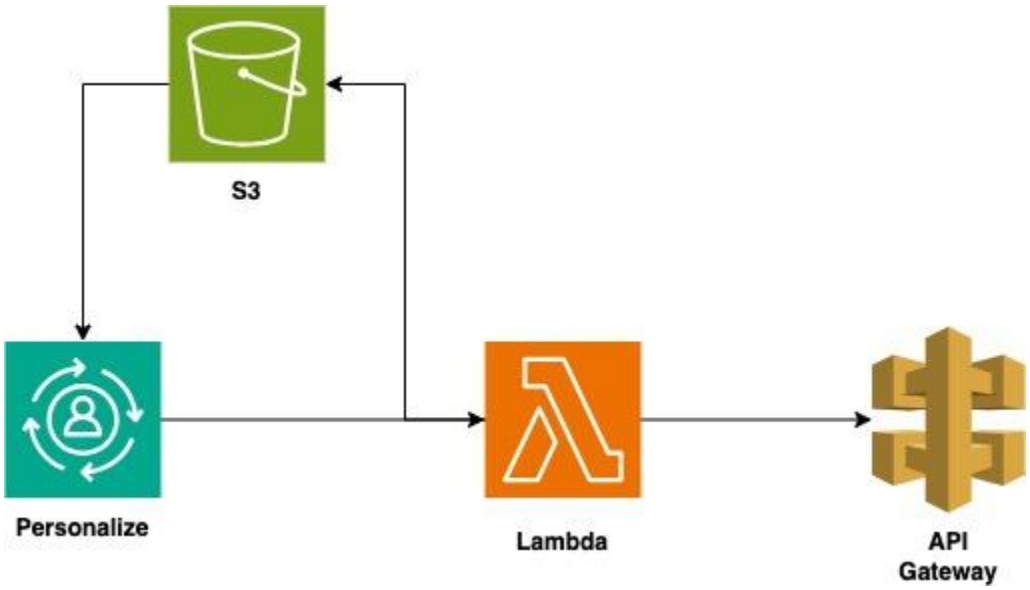
Video Recommendation System Documentation

Introduction:

The Video Recommendation System is a personalized recommendation engine built using AWS services like Amazon Personalize, Amazon S3, AWS Lambda, and API Gateway. The system aims to provide users with personalized video recommendations based on their past interactions with videos, such as watching duration and engagement.

Architecture Overview:



The architecture of the Video Recommendation System consists of several components:

Amazon Personalize: Amazon Personalize is a machine learning service provided by AWS that enables developers to create personalized recommendations for their applications. In our system, we use Amazon Personalize to train recommendation models based on user interactions with videos.

Amazon S3: Amazon S3 (Simple Storage Service) is a scalable object storage service provided by AWS. We use S3 to store the dataset containing user interactions with videos. Additionally, S3 is used to store the video files themselves.

AWS Lambda: AWS Lambda is a serverless compute service that allows running code without provisioning or managing servers. In our system, Lambda functions are used for preprocessing the dataset, retrieving video recommendations from Amazon Personalize, and serving recommendations via API Gateway.
API Gateway: API Gateway is a fully managed service that makes it easy for developers to create, publish, maintain, monitor, and secure APIs. We use API Gateway to expose an HTTP endpoint for accessing video recommendations generated by our system.

Data Preprocessing:

Before training the recommendation models, data preprocessing is performed to prepare the dataset for training. The main preprocessing steps include:

- **Deciding Whether a Video was Watched or Clicked:** This step involves determining whether a user's interaction with a video constitutes watching or just clicking based on the watching duration. For example, if a user watches more than 50% of a video's duration, it can be considered as watched, otherwise, it's considered as clicked.
- **Factorization Matrix:** The dataset is transformed into a factorization matrix where each row represents a user and each column represents an item (video). The matrix is then used as input for training the recommendation models in Amazon Personalize.

Recommendation Generation:

Once the recommendation models are trained using the preprocessed dataset, they are ready to generate personalized video recommendations. The recommendation generation process involves the following steps:

Training Models: Amazon Personalize trains multiple recommendation models based on the dataset provided. These models include user-personalization, item-personalization, and related algorithms.
Retrieving Recommendations: When a user makes a request for video recommendations through the API Gateway endpoint, a Lambda function is triggered. The Lambda function interacts with Amazon Personalize to retrieve personalized recommendations for the user based on their past interactions.
Formatting Recommendations: The retrieved recommendations are formatted into a user-friendly response, typically in JSON format. Each recommendation includes details such as the recommended video's ID, title, URL, and relevance score.

Returning Recommendations: The formatted recommendations are returned to the user through the API Gateway endpoint, where they can be consumed by client applications, such as web or mobile apps.

Dataset Preprocess: Preparing Dataset for AWS Personalize and Creating Dataset Groups

1. **Data Collection and Preprocessing:**
    a. **Data Collection:** Gather user interaction data with items and any relevant item metadata. This data can include user clicks, views, ratings, purchases, etc.

    b. **Data Preprocessing:** Clean the collected data and handle any missing or inconsistent values. Ensure the data is formatted correctly and encode categorical variables as needed.

2. **Dataset Creation:**
    a. **Login to AWS Console:** Access the AWS Management Console using your AWS account credentials.

    b. **Open AWS Personalize Console:** Navigate to the AWS Personalize service within the AWS Console.

    c. **Create Dataset Group:** Click on "Dataset groups" in the left navigation pane. Click on "Create dataset group" and provide a name for the group.

    d. **Create Dataset:** Within the dataset group, click on "Datasets" and then "Create dataset". Select the appropriate dataset type (Interactions, Items, Users) based on your data. Define the schema for the dataset according to your data structure. Upload your preprocessed data either from Amazon S3 or your local machine. Initiate the import process to import the data into the dataset.

3. **Dataset Group Creation:**
    a. **Open Dataset Groups**: Navigate to "Dataset groups" in the AWS Personalize console.

    b. Create Dataset Group: Click on "Create dataset group". Provide a name for the dataset group and optionally specify a role ARN for dataset import jobs. Proceed to the next step.

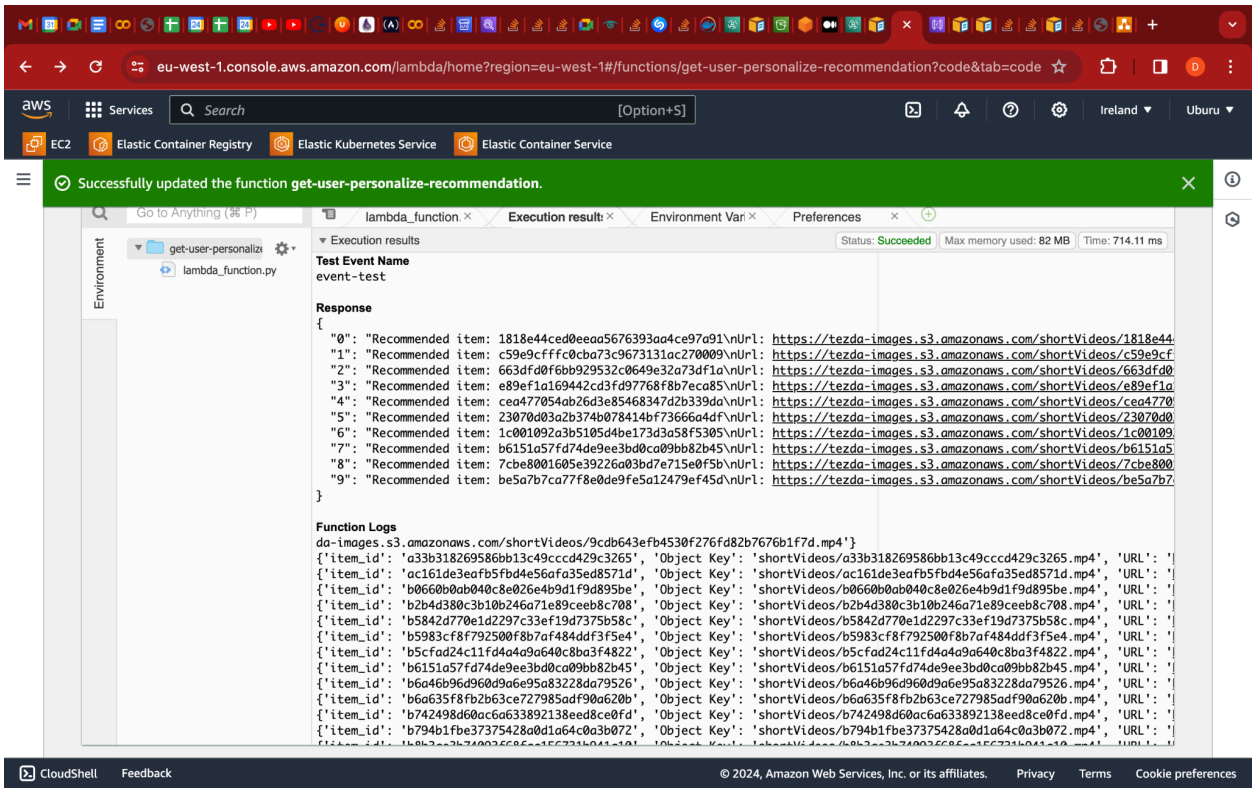
    c. Configure Dataset Group: Select the appropriate datasets for interactions, items, and users. Configure event types and event trackers if needed.

Work left to be done:

1. Integrating the endpoint into Streamlit for visualization.
2. Improving calibration by loading in more data.

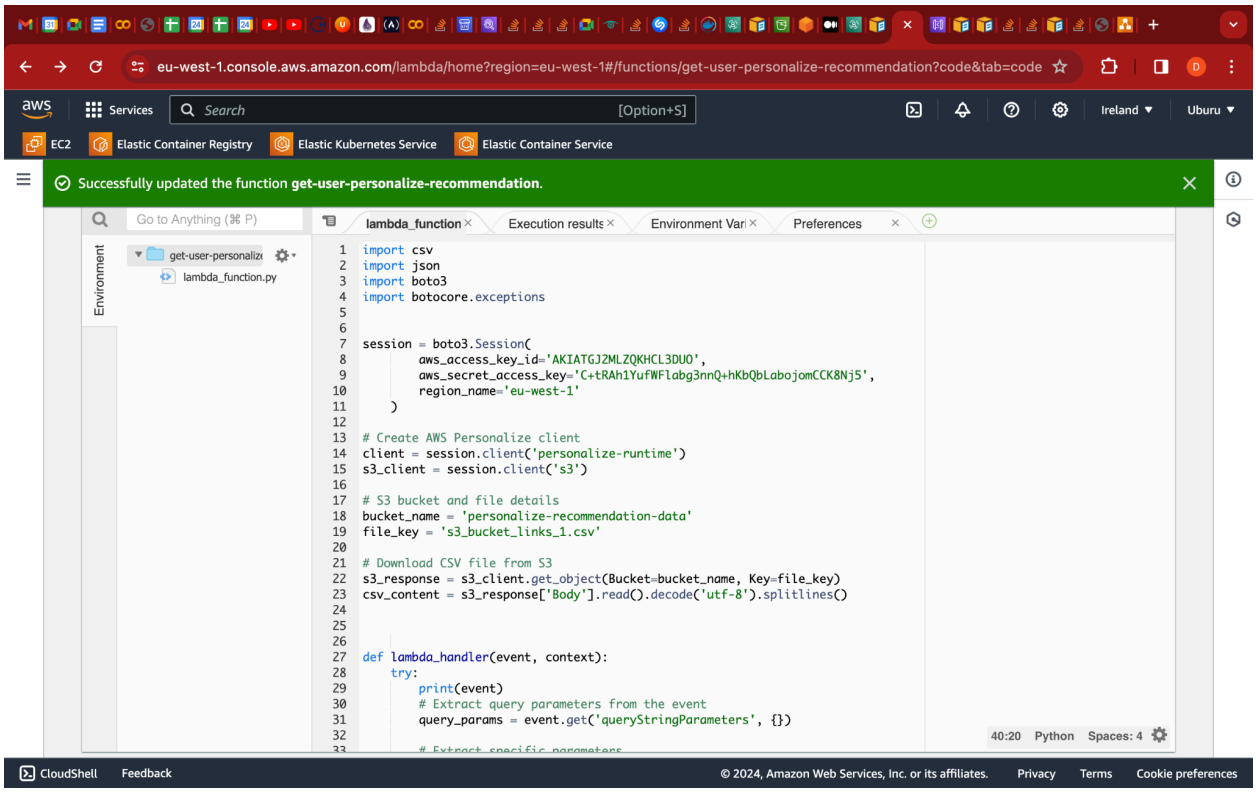
Screenshots:

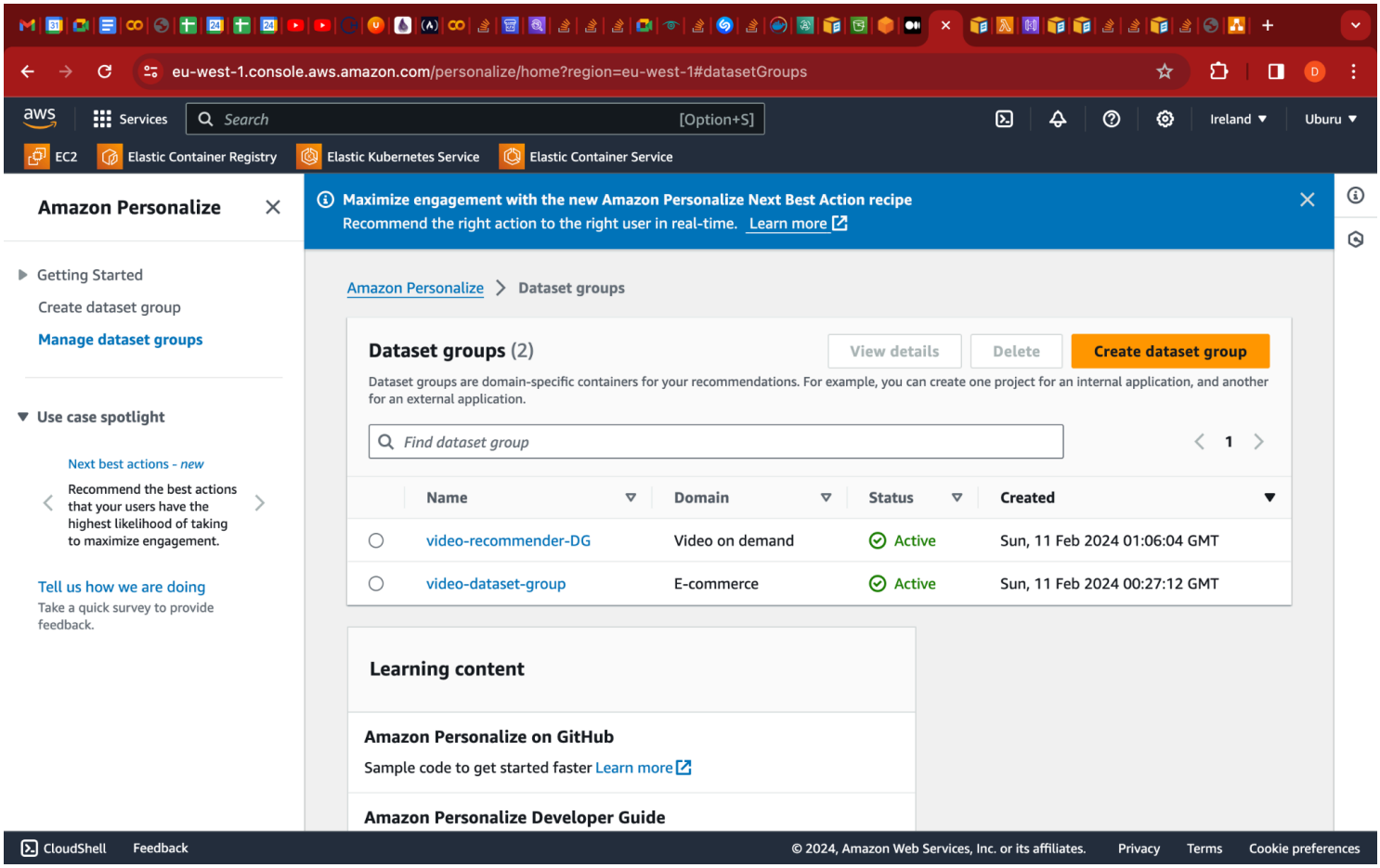
1. Test result of user recommendation from lambda:

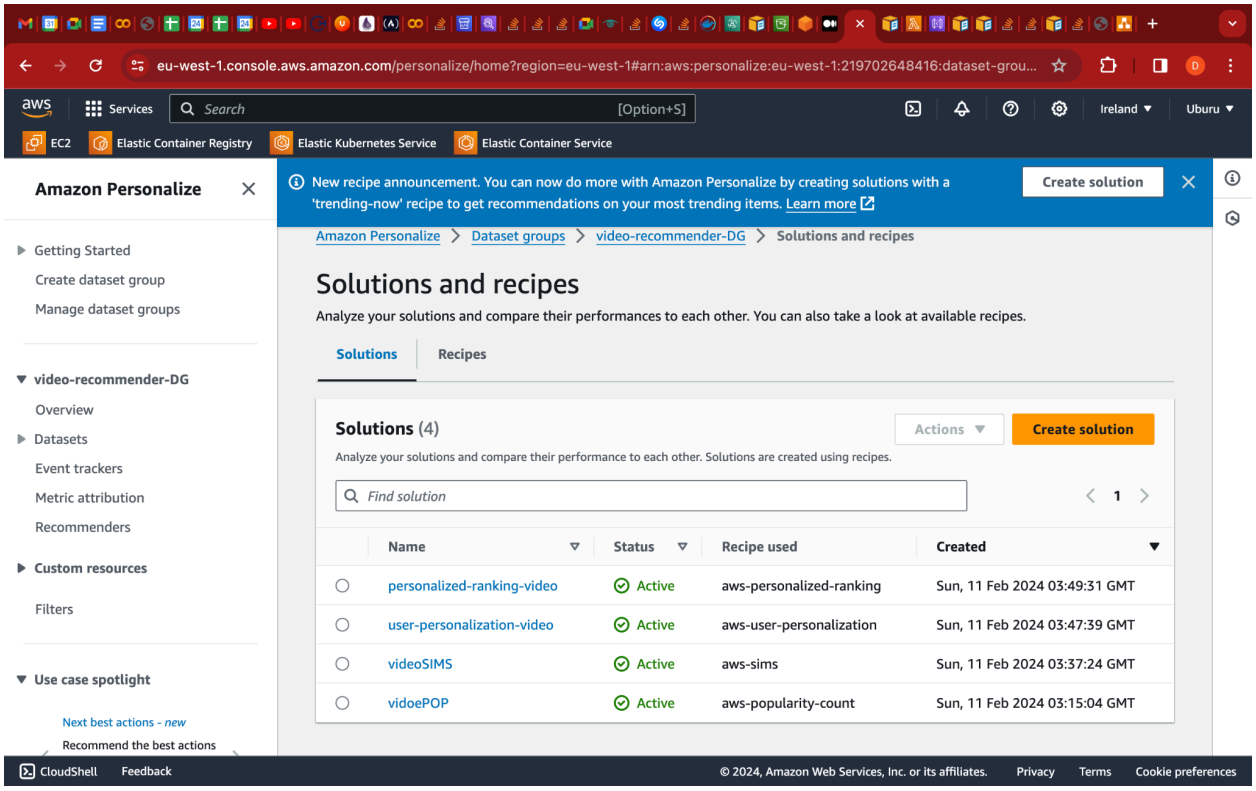


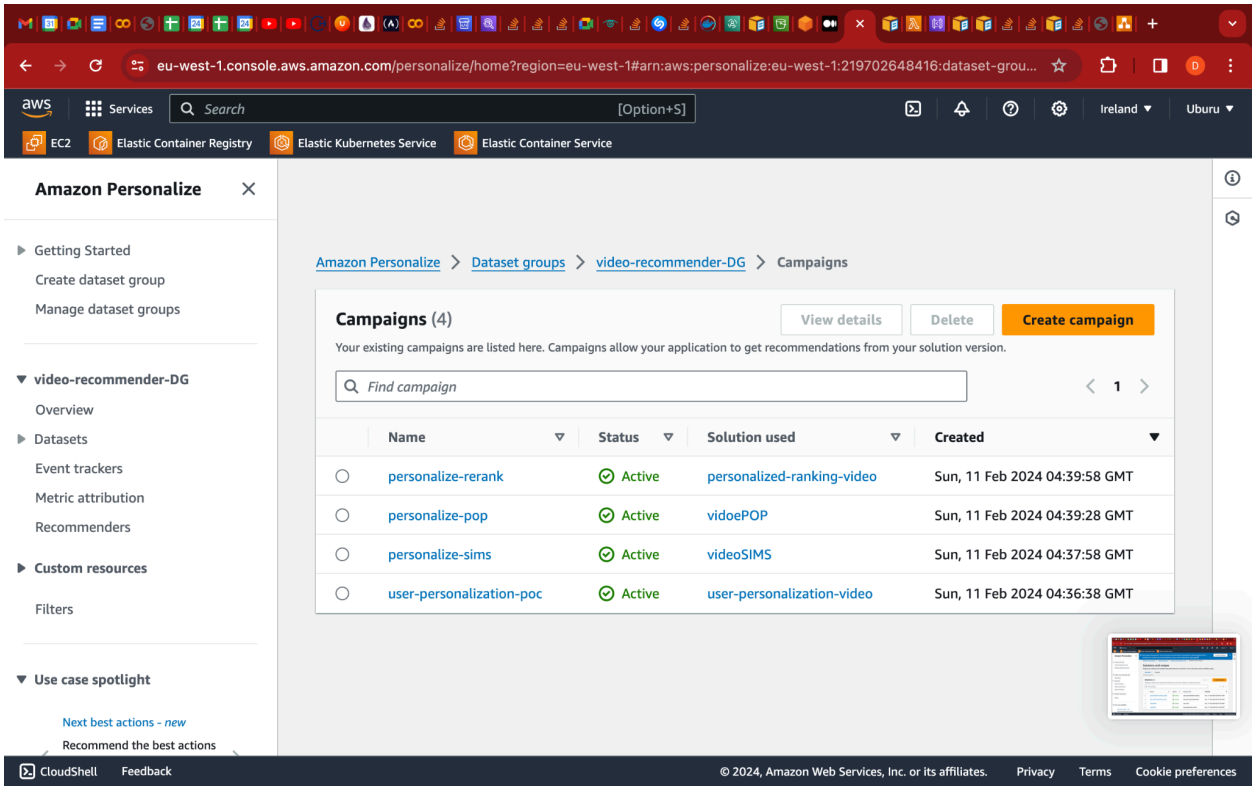
2

Lambda function code:






eu-west-1.console.aws.amazon.com/personalize/home?region=eu-west-1#arn:aws:personalize:eu-west-1:219702648416:dataset-grou...
Services
Search
[Option+S]
Ireland
Uburu
EC2
Elastic Container Registry
Elastic Kubernetes Service
Elastic Container Service
Amazon Personalize
New recipe announcement. You can now do more with Amazon Personalize by creating solutions with a 'trending-now' recipe to get recommendations on your most trending items. Learn more
Create solution
Getting Started
Create dataset group
Manage dataset groups
video-recommender-DG
Overview
Datasets
Event trackers
Metric attribution
Recommenders
Custom resources
Filters
Use case spotlight
Next best actions - new
Recommend the best actions
Amazon Personalize > Dataset groups > video-recommender-DG > Solutions and recipes
Solutions and recipes
Analyze your solutions and compare their performances to each other. You can also take a look at available recipes.
Solutions
Recipes
Solutions (4)
Actions
Create solution
Analyze your solutions and compare their performance to each other. Solutions are created using recipes.
Find solution
1
Name
Status
Recipe used
Created
personalized-ranking-video
Active
aws-personalized-ranking
Sun, 11 Feb 2024 03:49:31 GMT
user-personalization-video
Active
aws-user-personalization
Sun, 11 Feb 2024 03:47:39 GMT
videoSIMS
Active
aws-sims
Sun, 11 Feb 2024 03:37:24 GMT
vidoePOP
Active
aws-popularity-count
Sun, 11 Feb 2024 03:15:04 GMT
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences


eu-west-1.console.aws.amazon.com/personalize/home?region=eu-west-1#arn:aws:personalize:eu-west-1:219702648416:dataset-grou...
Services
Search
[Option+S]
Ireland
Uburu
EC2
Elastic Container Registry
Elastic Kubernetes Service
Elastic Container Service
Amazon Personalize
Getting Started
Create dataset group
Manage dataset groups
video-recommender-DG
Overview
Datasets
Event trackers
Metric attribution
Recommenders
Custom resources
Filters
Use case spotlight
Next best actions - new
Recommend the best actions
Amazon Personalize > Dataset groups > video-recommender-DG > Campaigns
Campaigns (4)
View details
Delete
Create campaign
Your existing campaigns are listed here. Campaigns allow your application to get recommendations from your solution version.
Find campaign
1
Name
Status
Solution used
Created
personalize-rerank
Active
personalized-ranking-video
Sun, 11 Feb 2024 04:39:58 GMT
personalize-pop
Active
vidoePOP
Sun, 11 Feb 2024 04:39:28 GMT
personalize-sims
Active
videoSIMS
Sun, 11 Feb 2024 04:37:58 GMT
user-personalization-poc
Active
user-personalization-video
Sun, 11 Feb 2024 04:36:38 GMT
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences