

TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT TP HCM
KHOA ĐÀO TẠO CHẤT LƯỢNG CAO



ĐỀ TÀI:
TÌM HIỂU VỀ DỊCH VỤ RECOMMENDATION CỦA
AWS

Giảng viên hướng dẫn: TS. Huỳnh Xuân Phụng

Nhóm sinh viên thực hiện:

Cao Xuân Thắng – 19119222

Trần Minh Sang – 19110278

Tp. Hồ Chí Minh, tháng 11 năm 2021

MỤC LỤC

MỤC LỤC	1
DANH MỤC HÌNH ẢNH	1
Phần 1: Phần mở đầu	2
1.1. Khái quát về Recommendation System	3
1.2. Mục tiêu của Recommendation System.....	3
1.3. Một số ví dụ về hệ thống gợi ý (Recommendation System)	4
1.4. Phương pháp tiếp cận.....	5
1.4.1. Collaborative Filtering	6
1.4.2. Content-Based Filtering	6
1.4.3. Hybrid Recommendation Systems	7
Phần 2: Dịch vụ Recommendation của AWS	8
1. Giới thiệu Recommendation của Amazon Personalize	8
2. Thuật toán triển khai	9
3. Đặc điểm nổi bật.....	11
4. Cách cài đặt.....	12
Phần 3: Phần kết.....	23
Phần 4: Tài liệu tham khảo	24

DANH MỤC HÌNH ẢNH

Hình 1 Create bucket	12
Hình 2 Bỏ chọn Block all public access.....	13
Hình 3 Sau khi tạo Bucket	14
Hình 4 Tải dữ liệu từ máy lên.....	15
Hình 5 Đã tạo xong dữ liệu Bucket.....	16
Hình 6 Dataset group	16
Hình 7 Sau khi tạo Dataset group	17
Hình 8 Import User-item	17
Hình 9 Tạo new schema	18
Hình 10 Nhập tên Dataset inport job name.....	19
Hình 11 Copy URL S3.....	19
Hình 12 Paste URL vào Data location.....	20
Hình 13 Kết quả sau khi tạo Schema.....	20
Hình 14 Vào Solution and recipes, chọn Create solution	21
Hình 15 Nhập đầy đủ thông tin và nhấn Next	21
Hình 16 Kết quả sau khi tạo solution	22
Hình 17 tại Campains chọn Create campaign	22
Hình 18 Nhập thông tin và chọn Create campaign	23

Phần 1: Phần mở đầu

1.1. Khái quát về Recommendation System

Với sự phát triển của công nghệ 4.0 và sự phát triển nhanh chóng của thương mại điện tử cùng với sự tương tác trực tiếp online của mọi người hiện nay và trong tương lai gần, hệ thống gợi ý (Recommendation systems) đã được phát triển và ứng dụng vào nhiều trang web hay dịch vụ mà chúng ta tương tác hằng ngày. Hệ thống gợi ý là một lớp con của hệ thống lọc thông tin mà dự đoán trước cho người sử dụng cần tìm dữ liệu gì. Được sử dụng rộng rãi trong nhiều lĩnh vực trong cuộc sống và được sử dụng phổ biến trong nhiều công ty thương mại điện tử hiện nay. Ý tưởng cho hệ thống gợi ý này là ta sử dụng các dữ liệu thu nhập dữ liệu từ người dùng hay lịch sử người dùng đã tìm kiếm trước đó và dự đoán dữ liệu mà người dùng sẽ tìm kiếm, gợi ý cho người dùng những sản phẩm, tính năng hay dịch vụ mà người dùng cần đến. Đối tượng mà hệ thống gợi ý hướng đến là user, còn sản phẩm hay dữ liệu hệ thống đưa ra gợi ý là item. Qua đó, chúng ta thấy được hệ thống gợi ý dựa trên tương tác trong quá khứ giữa user và item để dự đoán cho người dùng trong tương lai.

1.2. Mục tiêu của Recommendation System

Để giúp cho người dùng có thể tìm kiếm nhanh chóng, phù hợp với người dùng cũng như trong kinh doanh giúp cho tăng doanh thu và đem lại lợi nhuận cho mặt hàng đó, đem hướng chú ý của user về những item đó mà hệ thống gợi ý đã đưa ra một số mục tiêu và tiêu chí kỹ thuật:

- + Relevance (Tính phù hợp): Mục tiêu hiển nhiên nhất là gợi ý items thích hợp với users, do users thường sử dụng những items mà họ cảm thấy thú vị và phù hợp. Mặc dù sự thích hợp là mục tiêu chính của hệ thống gợi ý, nhưng nếu chỉ có mỗi mục tiêu này thì vẫn là chưa đủ, vì vậy chúng ta sẽ nói đến một số mục tiêu "hạng 2" sau đây, mặc dù nó không quan trọng bằng Tính phù hợp, nhưng nó vẫn mang một ý nghĩa quan trọng và ảnh hưởng khá nhiều đến một hệ thống gợi ý được xem là tốt.

- + Novelty (Tính mới lạ): Hệ thống gợi ý thực sự có hiệu quả khi nó gợi ý những items mà users chưa từng thấy trong quá khứ. Ví dụ, một bộ phim phổ biến với một thể loại quen thuộc sẽ hiếm khi gây ấn tượng cho user. Việc lặp đi lặp lại gợi ý những items phổ biến cũng có thể dẫn tới sự giảm sút trong sự đa dạng trong việc "bán" items.
- + Serendipity (tính tình cờ): Trong mục tiêu này, items được gợi ý là một items không ngờ tới đối với user, và vì vậy, đây có thể được coi là cách gợi ý mang yếu tố may mắn, trái ngược với các mục tiêu gợi ý trước đó. Serendipity khác với Novelty ở chỗ những gợi ý này thực sự gây ngạc nhiên với user, thay vì gợi ý những items mà user chưa thấy trước đó. Mục tiêu này có thể được sử dụng trong trường hợp mà user đó chỉ quen thuộc với một loại items cụ thể, mà không biết đến những loại items khác mà có thể gây ngạc nhiên và thích thú đối với những loại items mới này. Không giống với Novelty, Serendipity tìm cách mở rộng sự gợi ý đối với user. Đây có thể coi là một chiến lược mang tính lâu dài, nhằm mang đến những sở thích mới, những mối quan tâm mới cho users, từ đó có thể nâng cao được doanh số và lợi nhuận.
- + Mở rộng sự đa dạng trong việc gợi ý: Hệ thống gợi ý thường gợi ý một danh sách top-k items. Nếu chúng quá giống nhau, có thể xảy ra trường hợp mà user có thể sẽ không thích items nào cả, và việc đó hoàn toàn không có lợi. Mặt khác, khi hệ thống gợi ý cố gắng đưa ra gợi ý các items ở nhiều loại khác nhau, khả năng mà user thích ít nhất một item sẽ tăng lên đáng kể, giảm được rủi ro trong việc gây nhàm chán cho user với việc gợi ý lặp đi lặp lại những sản phẩm quá giống nhau.

1.3. Một số ví dụ về hệ thống gợi ý (Recommendation System)

GroupLens Recommender System

GroupLens được coi là hệ thống gợi ý tiên phong, nó được xây dựng để gợi ý cho Usenet news. Hệ thống sẽ thu thập những ratings của người đọc Usenet và sử dụng chúng để dự đoán liệu một người dùng có thích bài báo hay không trước khi họ đọc chúng. Trong hệ thống này, một số thuật toán "collaborative filtering" đã được sử

dụng. Một số ý tưởng sau đó cũng được phát triển dựa theo hệ thống này, những ý tưởng đó áp dụng cho sách và phim, với bộ dữ liệu khá quen thuộc cho những người mới bắt đầu là BookLens và MovieLens.

Amazon Recommender System

Amazon cũng được coi là một trong những hệ thống gợi ý tiên phong, đặc biệt là trong thương mại và hệ thống gợi ý đã được sử dụng rất thành công ở Amazon. Hệ thống này sử dụng đánh giá của người dùng, hành vi của người dùng và lịch sử hoạt động của người dùng để làm dữ liệu dự đoán. Ratings ở Amazon có thang điểm từ 1 đến 5, còn lịch sử mua hàng và lịch sử hoạt động có thể dễ dàng được thu thập bằng khi mà user thực hiện đăng nhập trên trang web của Amazon. Những gợi ý cho người dùng được hiển thị trên trang chủ của Amazon, ngay khi người dùng truy cập vào trang web. Trong nhiều trường hợp, những items được gợi ý cũng được giải thích, ví dụ như trong mối quan hệ của sản phẩm được gợi ý với những sản phẩm mà user đã mua trước đó.

Facebook Friend Recommendations

Facebook thường gợi ý cho chúng ta kết bạn với những người mà chúng ta có thể quen, nhằm mục đích tăng số lượng kết nối trên mạng xã hội. Kiểu gợi ý này với mục tiêu có chút khác biệt so với hệ thống gợi ý sản phẩm. Trong khi hệ thống gợi ý sản phẩm trực tiếp tạo ra lợi nhuận cho người cung cấp dịch vụ bằng cách tăng doanh số sản phẩm bán được, thì trong mạng xã hội, việc gợi ý kết bạn nhằm tăng số lượng kết nối trên mạng xã hội và từ đó nâng cao trải nghiệm của người dùng. Cùng với việc tăng được số lượng kết nối, lợi nhuận thu được từ việc quảng cáo cũng sẽ tăng lên.

1.4. Phương pháp tiếp cận

Các hệ thống khuyến nghị/ gợi ý (Recommendation System/ Recommender) được xây dựng dựa trên 3 hướng tiếp cận đó là Collaborative Filtering (CF), Content-Based Filtering và Hybrid Recommendation Systems.

1.4.1. Collaborative Filtering

Collaborative Filtering (Lọc cộng tác) dựa trên item đã tìm kiếm trong quá khứ sẽ tìm kiếm trong tương lai và họ sẽ thích những item tương tự. Hệ thống chỉ tạo các đề xuất bằng cách sử dụng thông tin về cấu hình xếp hạng cho những người dùng hoặc mặt hàng khác nhau. Bằng cách định vị người dùng/ item ngang hàng có lịch sử xếp hạng tương tự với người dùng hoặc mặt hàng hiện tại, họ tạo các đề xuất bằng cách sử dụng vùng lân cận này. Các phương pháp lọc cộng tác được phân loại là dựa trên bộ nhớ và dựa trên mô hình.

Khi xây dựng một mô hình từ hành vi của người dùng, người ta thường phân biệt giữa các hình thức thu thập dữ liệu rõ ràng và ngầm định.

Các phương pháp lọc cộng tác thường gặp phải ba vấn đề: cold start, scalability và sparsity.

Bắt đầu lạnh (Cold start): Đối với người dùng hoặc mặt hàng mới, không có đủ dữ liệu để đưa ra các đề xuất chính xác.

Khả năng mở rộng (Scalability): Trong nhiều môi trường mà các hệ thống này đưa ra đề xuất, có hàng triệu người dùng và sản phẩm. Do đó, một lượng lớn sức mạnh tính toán thường là cần thiết để tính toán các khuyến nghị.

Độ thưa thớt (Sparsity): Số lượng item được tìm kiếm trên các trang thương mại điện tử lớn là vô cùng lớn. Những người dùng tích cực nhất sẽ chỉ đánh giá một tập hợp con nhỏ của cơ sở dữ liệu tổng thể. Vì vậy, ngay cả những mặt hàng phổ biến nhất cũng có rất ít xếp hạng.

1.4.2. Content-Based Filtering

Content-Based Filtering (Phương pháp lọc dựa trên nội dung) dựa trên mô tả của mặt hàng và hồ sơ về sở thích của người dùng. Phương pháp này phù hợp nhất với các trường hợp có dữ liệu đã biết về một mục (tên, vị trí, mô tả, v.v.), nhưng trong user không có dữ liệu.

Các từ khóa được sử dụng mô tả các mặt hàng và hồ sơ người dùng xây dựng chỉ ra loại mặt hàng mà người dùng thích. Có thể nói là đề xuất các mục tương

tự như những mục mà người dùng đã thích trong quá khứ hoặc đang kiểm tra ở hiện tại.

Để tạo hồ sơ người dùng, hệ thống chủ yếu tập trung vào hai loại thông tin:

1. Một mô hình sở thích của người dùng.
2. Lịch sử tương tác của người dùng với hệ thống giới thiệu.

Cơ bản, phương pháp này sử dụng một hồ sơ item (một tập hợp các thuộc tính và tính năng rời rạc) đặc trưng cho item trong hệ thống. Hệ thống tạo hồ sơ dựa trên nội dung của người dùng dựa trên vectơ trọng số của các tính năng của mặt hàng.

Hệ thống gợi ý dựa trên nội dung cũng có thể bao gồm hệ thống tư vấn dựa trên ý kiến. Trong một số trường hợp, người dùng được phép đánh giá hay phản hồi về item. Các văn bản đánh giá do người dùng tạo là dữ liệu ngầm cho hệ thống giới thiệu. Các phương pháp tiếp cận phổ biến của hệ thống tư vấn dựa trên ý kiến sử dụng các kỹ thuật khác nhau bao gồm khai thác văn bản, truy xuất thông tin, phân tích tình cảm và học sâu.

1.4.3. Hybrid Recommendation Systems

Hầu hết các hệ thống gợi ý hiện nay sử dụng phương pháp kết hợp, kết hợp lọc cộng tác, lọc dựa trên nội dung và các phương pháp tiếp cận khác. Phương pháp tiếp cận kết hợp thực hiện bằng cách đưa ra các dự đoán dựa trên nội dung và dựa trên cộng tác một cách riêng biệt và sau đó kết hợp chúng; hay có thể thực hiện bằng cách thêm các khả năng dựa trên nội dung vào cách tiếp cận dựa trên cộng tác (và ngược lại); hoặc bằng cách thống nhất các phương pháp tiếp cận thành một mô hình. Các phương pháp này có thể được sử dụng khắc phục một số vấn đề phổ biến trong hệ thống khuyến nghị như khởi động nguội, thừa thớt.

Một số kỹ thuật lai bao gồm:

Weighted (Có trọng số): Kết hợp số điểm của các thành phần đề xuất khác nhau.

Switching (Chuyển đổi): Lựa chọn giữa các thành phần đề xuất và áp dụng thành phần đã chọn.

Mixed (Hỗn hợp): Các đề xuất từ những người giới thiệu khác nhau được trình bày cùng nhau để đưa ra đề xuất.

Feature Combination (Kết hợp tính năng): Các tính năng bắt nguồn từ các nguồn kiến thức khác nhau được kết hợp với nhau và đưa ra cho một thuật toán đề xuất duy nhất.

Feature Augmentation (Tăng cường tính năng): Tính toán một tính năng hoặc tập hợp các tính năng, sau đó là một phần của đầu vào cho kỹ thuật tiếp theo.

Cascade (Xếp tầng): Các đề xuất được ưu tiên nghiêm ngặt, với những đề xuất có mức độ ưu tiên thấp hơn sẽ phá vỡ mối quan hệ trong việc chấm điểm những đề xuất cao hơn.

Meta-level (Siêu cấp): Một kỹ thuật đề xuất được áp dụng và tạo ra một số loại mô hình, sau đó là đầu vào được sử dụng bởi kỹ thuật tiếp theo.

Tóm lại Collaborative Filtering là hướng tiếp cận mà chỉ dựa vào dữ liệu về hành vi tiêu dùng của người sử dụng trong khi Content-Based Filtering lại chỉ dựa vào các đặc điểm / đặc trưng của sản phẩm. Hybrid Recommendation Systems, thì lại dựa vào đồng thời cả dữ liệu về hành vi của người tiêu dùng và các đặc điểm của sản phẩm. Một ví dụ là hệ thống khuyến nghị của Netflix được sử dụng để khuyến nghị các bộ phim mà người dùng nên xem. Collaborative Filtering lại có hai hướng tiếp cận riêng biệt là User-base Collaborative Filtering và Item-base Collaborative Filtering.

Phần 2: Dịch vụ Recommendation của AWS

1. Giới thiệu Recommendation của Amazon Personalize

Amazon Personalize cho phép các nhà phát triển xây dựng ứng dụng bằng chính công nghệ máy học (Machine Learning) mà Amazon.com dùng để đề xuất các nội dung cá nhân hóa theo thời gian thực, kể cả khi họ chưa có chuyên môn về Machine Learning.

Amazon Personalize giúp dễ dàng xây dựng ứng dụng mang lại trải nghiệm cá nhân hóa, bao gồm cả nội dung đề xuất cho sản phẩm cụ thể, xếp hạng lại sản phẩm cá nhân hóa và tiếp thị trực tiếp tùy chỉnh. Amazon Personalize là một dịch vụ máy học được quản lý toàn phần. Dịch vụ này không sử dụng các hệ thống đề xuất dựa trên quy tắc cố định cứng nhắc mà sẽ đào tạo, tinh chỉnh và triển khai các mô hình Machine Learning tùy chỉnh để cung cấp các đề xuất được tùy chỉnh chi tiết cho khách hàng trong các ngành như bán lẻ, truyền thông và giải trí.

Amazon Personalize cung cấp cơ sở hạ tầng cần thiết và quản lý toàn bộ quy trình Machine Learning, bao gồm xử lý dữ liệu, xác định tính năng, sử dụng các thuật toán tốt nhất, đồng thời đào tạo, tối ưu hóa và lưu trữ các mô hình. Bạn sẽ nhận được kết quả thông qua Giao diện lập trình ứng dụng (API) và chỉ phải trả tiền cho những gì mình sử dụng, không cần nộp phí tối thiểu hoặc cam kết trả trước. Toàn bộ dữ liệu được mã hóa để đảm bảo tính bảo mật và riêng tư. Dữ liệu chỉ được dùng để tạo các đề xuất cho người dùng của bạn.

Nhờ có Amazon Personalize, có thể thiết kế và khởi chạy một công cụ đề xuất cá nhân. Công cụ đề xuất cá nhân dựa vào các dữ liệu đã cung cấp và hoạt động mà người dùng thực hiện thường xuyên trong quá khứ mà tạo thành một danh sách gợi ý (đề xuất) giúp cho người dùng thuận tiện hơn khi sử dụng và giúp hoàn thành công việc một cách nhanh chóng.

2. Thuật toán triển khai

Để triển khai dịch vụ Recommendation của AWS sau khi chuẩn bị dữ liệu đề xuất, ta nên chú ý đến một số thuật toán triển khai sau:

User-Personalization recipe

Công thức User-Personalization (aws-user-Personalization) được tối ưu hóa cho tất cả các trường hợp đề xuất được cá nhân hóa. Nó dự đoán các mục mà người dùng sẽ tương tác dựa trên các tập dữ liệu Interactions, Items, and User datasets. Khi đề xuất các mặt hàng, nó sử dụng tính năng tự động thăm dò Items.

Với tính năng khám phá tự động, Amazon Personalize tự động kiểm tra các đề xuất Items khác nhau, học hỏi từ cách người dùng tương tác với Items được đề xuất này

và tăng cường đề xuất cho Items thúc đẩy mức độ tương tác và chuyển đổi tốt hơn. Điều này cải thiện khả năng khám phá mặt hàng và mức độ tương tác khi bạn có một danh mục thay đổi nhanh hoặc khi Items mới, chẳng hạn như các bài báo hoặc chương trình khuyến mãi, phù hợp hơn với người dùng khi mới.

Bạn có thể cân bằng số lượng để khám phá (nơi các mục có ít dữ liệu tương tác hơn hoặc mức độ liên quan được đề xuất thường xuyên hơn) với số lượng cần khai thác (nơi các đề xuất dựa trên những gì chúng tôi biết hoặc mức độ liên quan). Amazon Personalize tự động điều chỉnh các đề xuất trong tương lai dựa trên phản hồi ngầm của người dùng.

Công thức được xác định trước này có các thuộc tính sau:

Name – aws-user-personalization

Recipe Amazon Resource Name (ARN) – arn:aws:personalize:::recipe/aws-user-personalization

Algorithm ARN – arn:aws:personalize:::algorithm/aws-user-personalization

Popularity-Count recipe

Mức độ phổ biến-Số lượng đề xuất Items phổ biến nhất dựa trên tất cả dữ liệu hành vi người dùng của bạn. Các Items phổ biến nhất có nhiều tương tác nhất với những người dùng duy nhất. Công thức trả về các mặt hàng phổ biến giống nhau cho tất cả người dùng. Popularity-Count là một cơ sở tốt để so sánh với các công thức khác bằng cách sử dụng các chỉ số đánh giá mà Amazon Personalize tạo ra khi bạn tạo một phiên bản giải pháp.

Công thức được xác định trước này có các thuộc tính sau:

Name – aws-popularity-count

Recipe ARN – arn:aws:personalize:::recipe/aws-popularity-count

Algorithm ARN – arn:aws:personalize:::algorithm/aws-popularity-count

Feature transformation ARN – arn:aws:personalize:::feature-transformation/sims

Recipe type – USER_PERSONALIZATION

Popularity-Count không có siêu tham số được hiển thị.

Personalized-Ranking recipe

Personalized-Ranking recipe tạo ra thứ hạng được cá nhân hóa của các mặt hàng. Một Personalized-Ranking là một danh sách các mục khuyến cáo rằng đang tái xếp cho một người dùng cụ thể. Điều này rất hữu ích nếu bạn có một bộ sưu tập các mục được sắp xếp, chẳng hạn như kết quả tìm kiếm, khuyến mại hoặc danh sách được sắp xếp và bạn muốn cung cấp một personalized re-ranking cho từng người dùng của mình.

Để đào tạo một mô hình, Personalized-Ranking recipe sử dụng dữ liệu trong tập dữ liệu Tương tác của bạn và nếu bạn đã tạo chúng, tập dữ liệu Mục và tập dữ liệu Người dùng trong nhóm tập dữ liệu của bạn. Với Personalized-Ranking, tập dữ liệu Items của bạn có thể bao gồm Unstructured text metadata và tập dữ liệu Interactions của bạn có thể bao gồm Contextual metadata.

Công thức được xác định trước này có các thuộc tính sau:

Name – aws-personalized-ranking

Recipe Amazon Resource Name (ARN) – arn:aws:personalize:::recipe/aws-personalized-ranking

Algorithm ARN – arn:aws:personalize:::algorithm/aws-personalized-ranking

Feature transformation ARN – arn:aws:personalize:::feature-transformation/JSON-percentile-filtering

Recipe type – PERSONALIZED_RANKING

3. Đặc điểm nổi bật

Cung cấp đề xuất chất lượng cao, theo thời gian thực: dùng thật toán Machine Learning để tạo ra các đề xuất chất lượng cao hơn đáp ứng được yêu cầu của người dùng.

Dễ dàng triển khai các đề xuất cá nhân hoá trong vài ngày: triển khai đề xuất cá nhân hoá tùy chỉnh được Machine Learning hỗ trợ giúp ta nhanh chóng tạo một cách dễ dàng với vài cú nhấp chuột

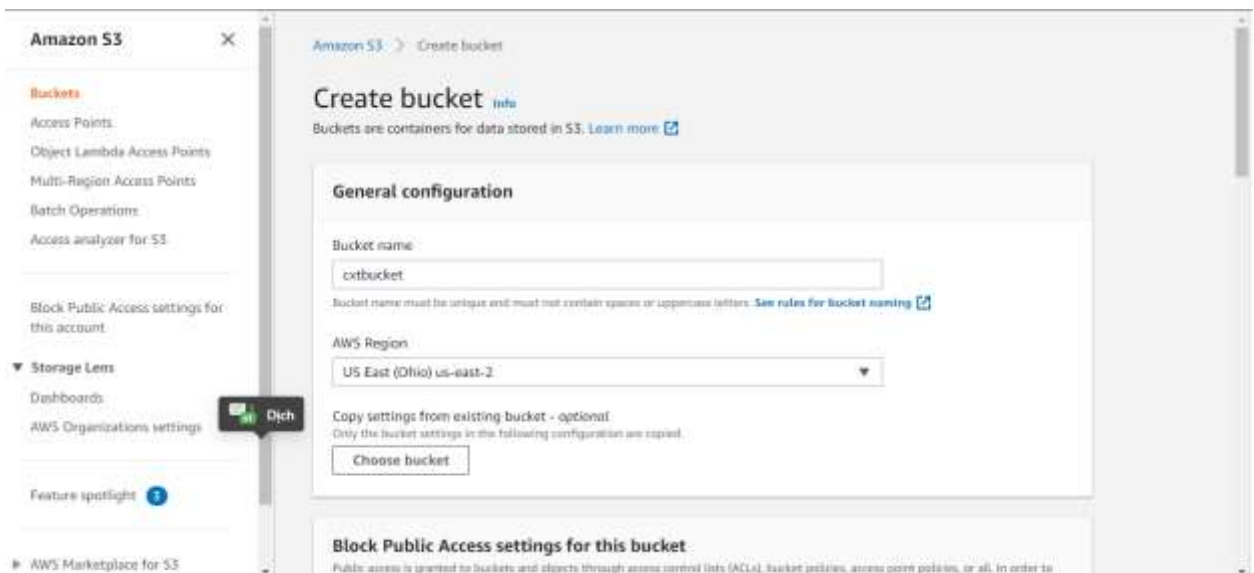
Bảo mật và quyền riêng tư về dữ liệu: toàn bộ dữ liệu được mã hoá, bảo mật, chỉ tạo đề xuất cho khách hàng của bạn, dữ liệu này không chia sẻ giữa khách hàng hay với Amazon.com.

Cá nhân hoá mọi điểm tiếp xúc trong quá trình của khách hàng: có thể thực hiện trên các kênh và thiết bị khác nhau mà không chi trả các chi phí cơ sở hạ tầng hoặc tài nguyên nào, cho phép sử dụng các đề xuất một cách linh hoạt trên thời gian thực dựa trên những yêu cầu của người dùng.

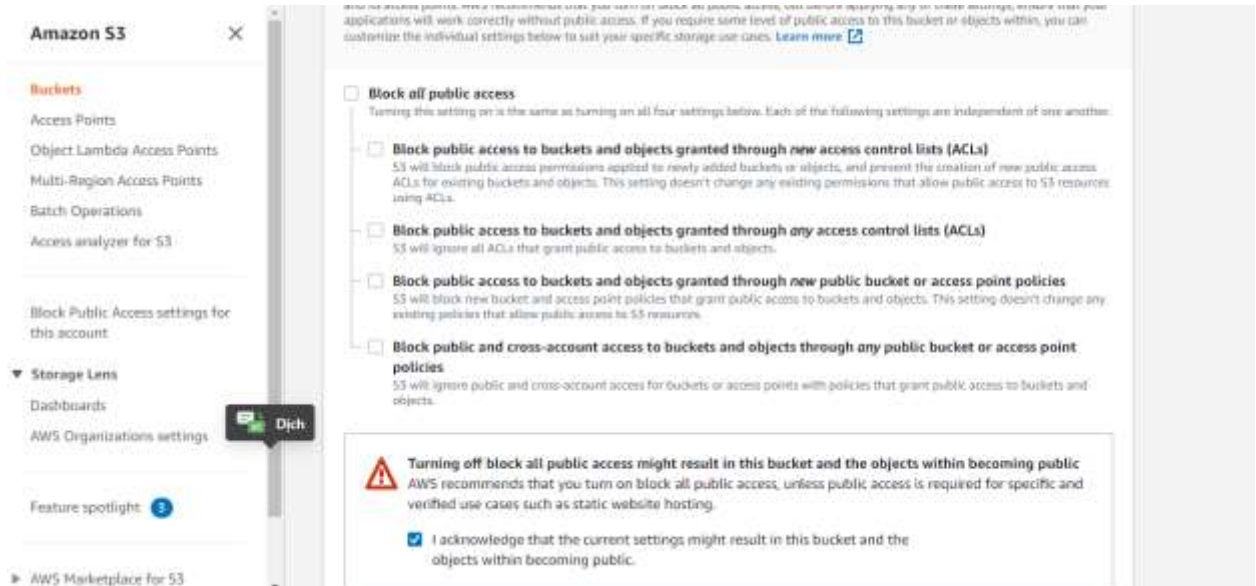
4. Cách cài đặt

Đầu tiên chúng ta phải tạo Buckets tại trang Amazon S3 của mình.

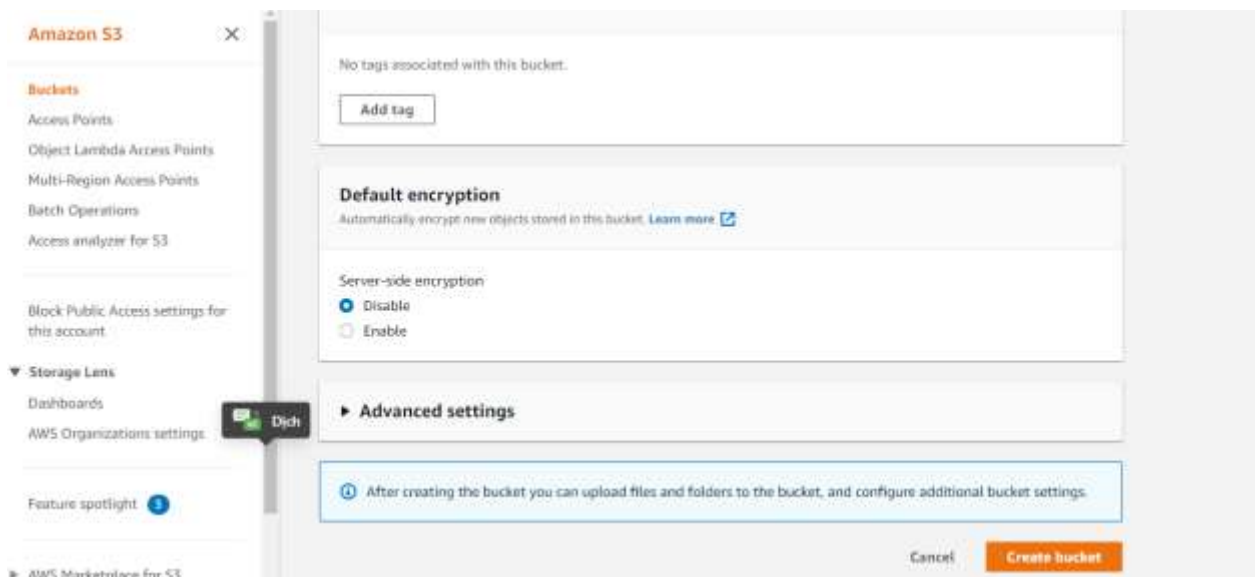
Ta bấm Create bucket -> Nhập Bucket name, Region -> Bỏ chọn Block all public access -> Phần Manage system permissions chọn Grant Amazon S3 Log Delivery group with access to this bucket -> Bấm Next -> Bấm Create bucket.

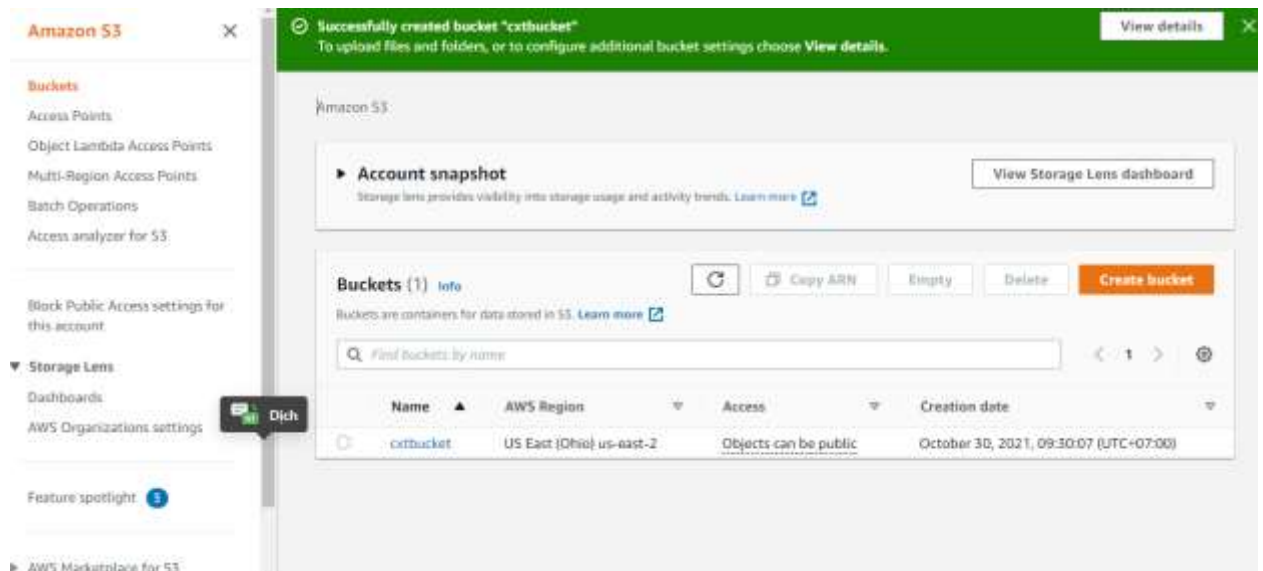


Hình 1 Create bucket



Hình 2 Bỏ chọn Block all public access





Hình 3 Sau khi tạo Bucket

Ta đã có Bucket. Chọn Bucket vừa tạo -> Tải dữ liệu phù hợp lên từ máy (.csv) -> Trong mục Permissions -> Ta vào Bucket Policy nhập vào phần Bucket policy editor đoạn code:

Bucket Policy json:

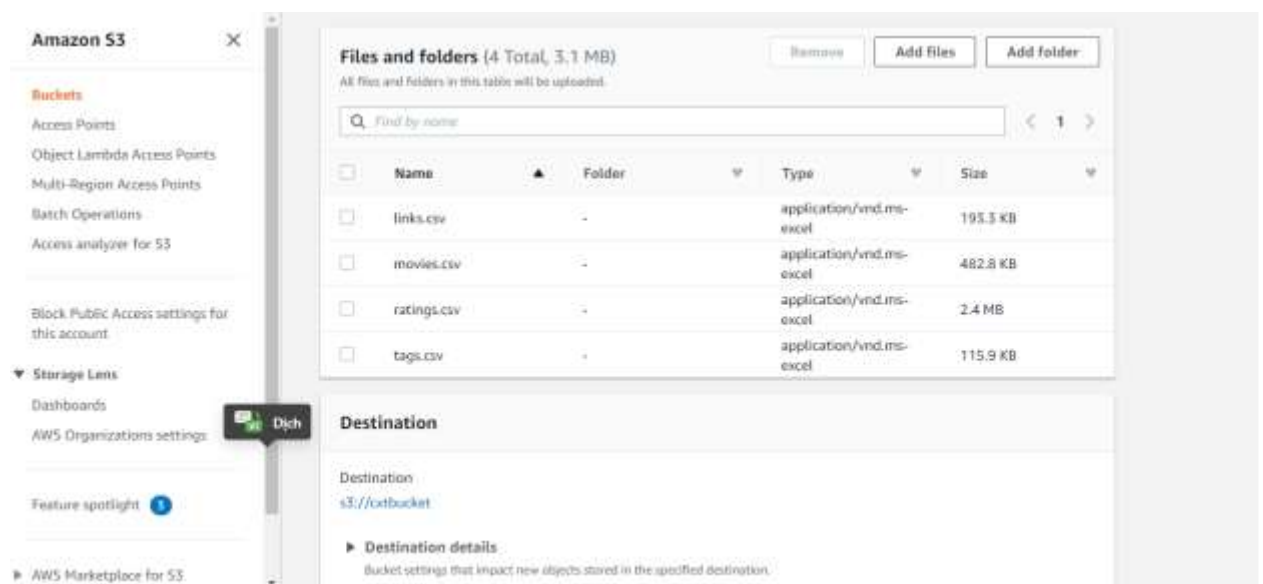
```
{
  "Version": "2012-10-17",
  "Id": "PersonalizeS3BucketAccessPolicy",
  "Statement": [
    {
      "Sid": "PersonalizeS3BucketAccessPolicy",
      "Principal": "*",
      "Effect": "Allow",
      "Action": [
        "s3:GetObject",
        "s3:ListBucket"
      ],

```

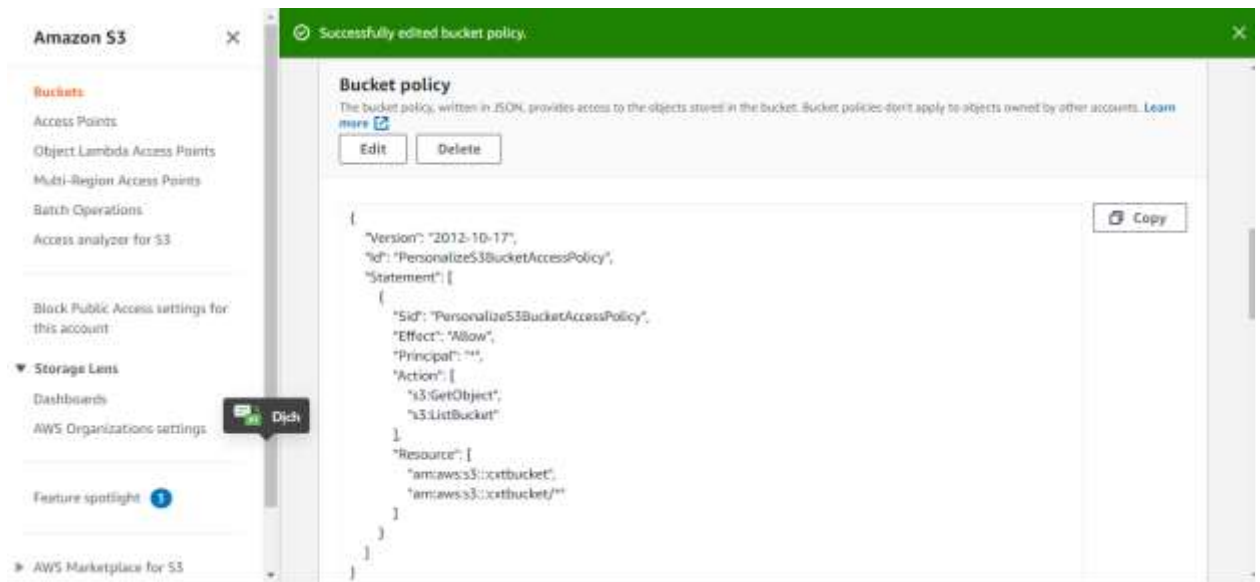
```

"Resource":
[
  "arn:aws:s3:::bucket-name",
  "arn:aws:s3:::bucket-name/*"
]
}
]
}

```

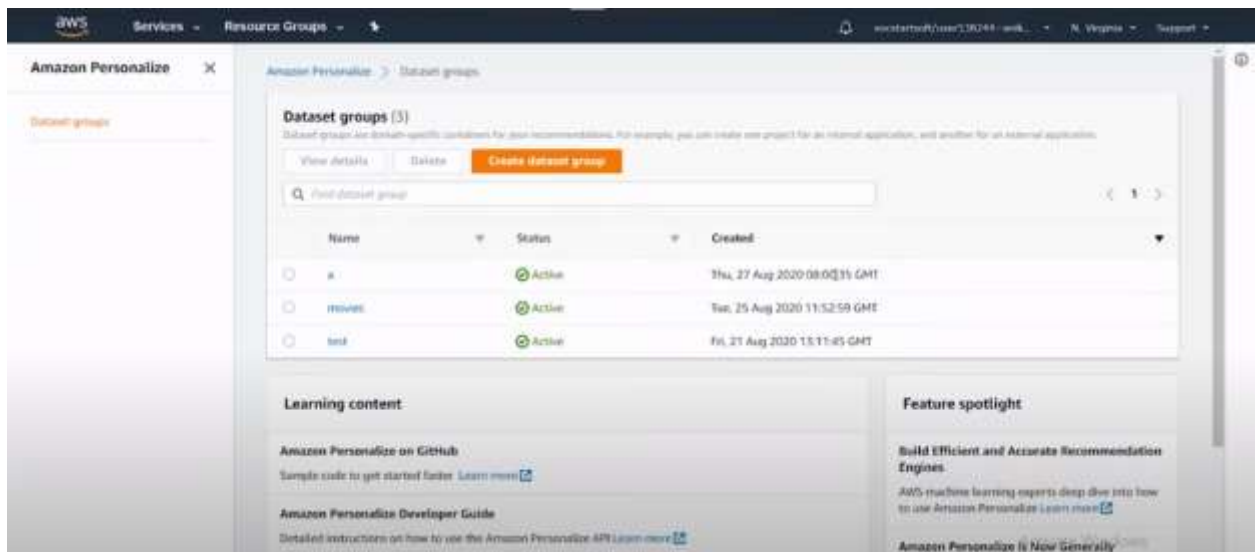


Hình 4 Tải dữ liệu từ máy lên

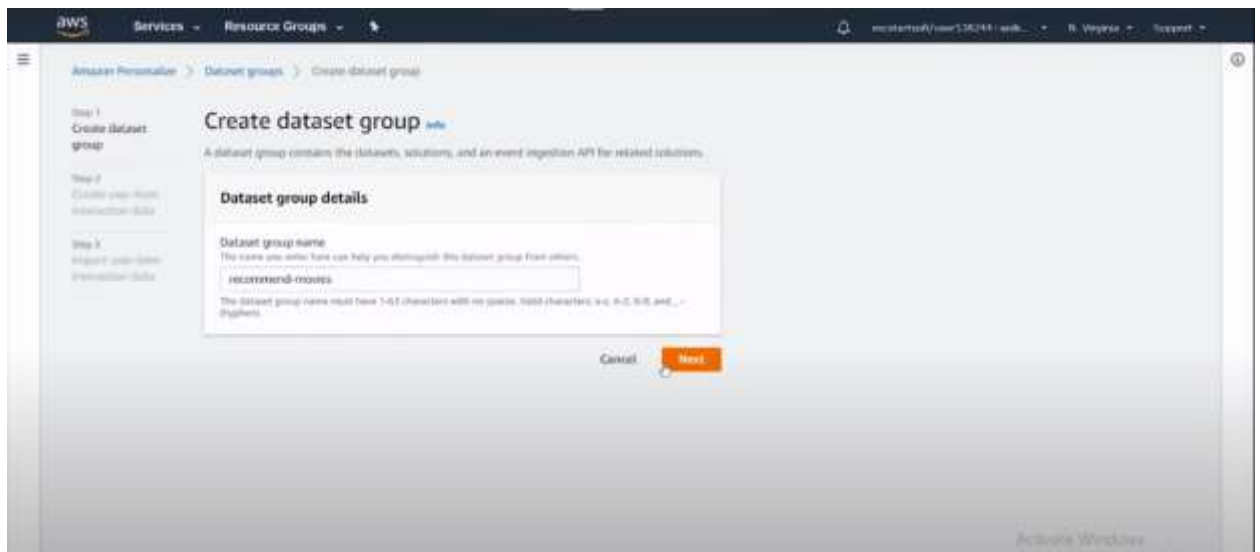


Hình 5 Đã tạo xong dữ liệu Bucket

Truy cập vào Amazon Personalize của mình -> Bấm vào Create dataset group -> Nhập tên cho Dataset group name -> Bấm Next.

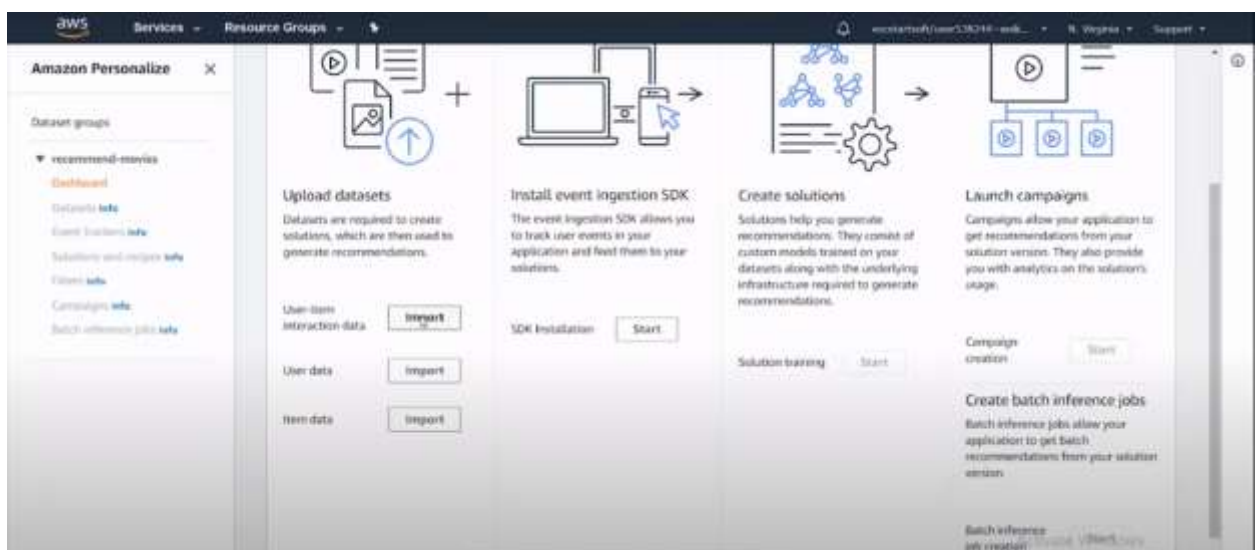


Hình 6 Dataset group



Hình 7 Sau khi tạo Dataset group

Quay lại trang Amazon Personalize của mình, bấm vào Dataset group vừa tạo -
 >Trong mục Dashboard -> Tại phần User-item interaction data -> Bấm Import.



Hình 8 Import User-item

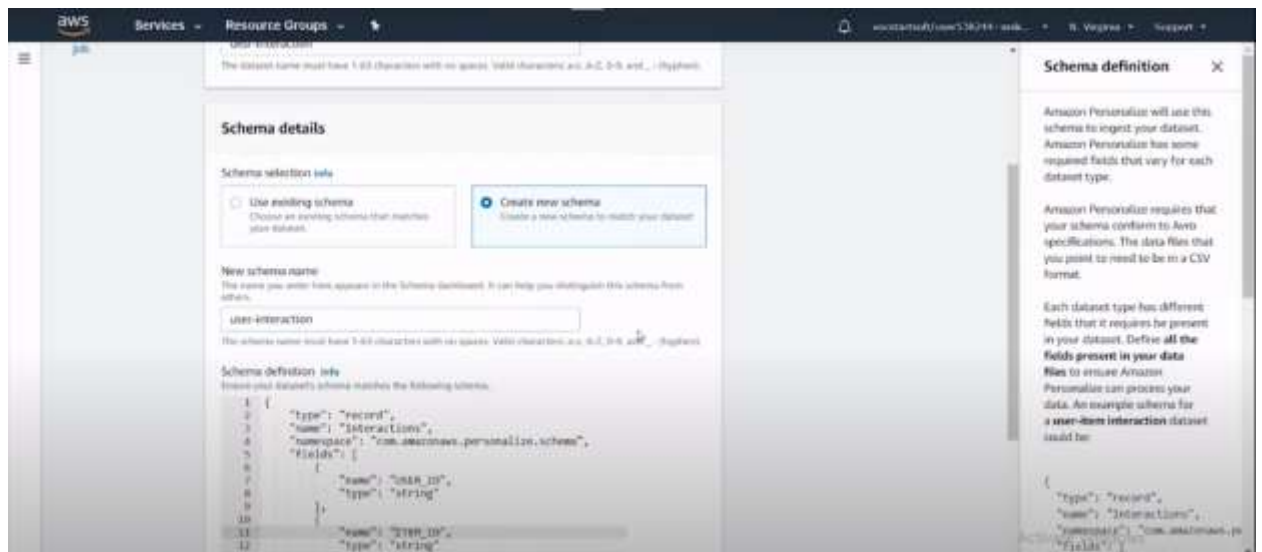
Nhập tên cho Dataset name -> Chọn vào Create new schema -> Nhập tên cho New schema name. Bởi vì data có dạng .csv nên bạn phải thay đổi Schema definition với định dạng:

```
{
  "type": "record",
  "name": "Interaction",
```

```

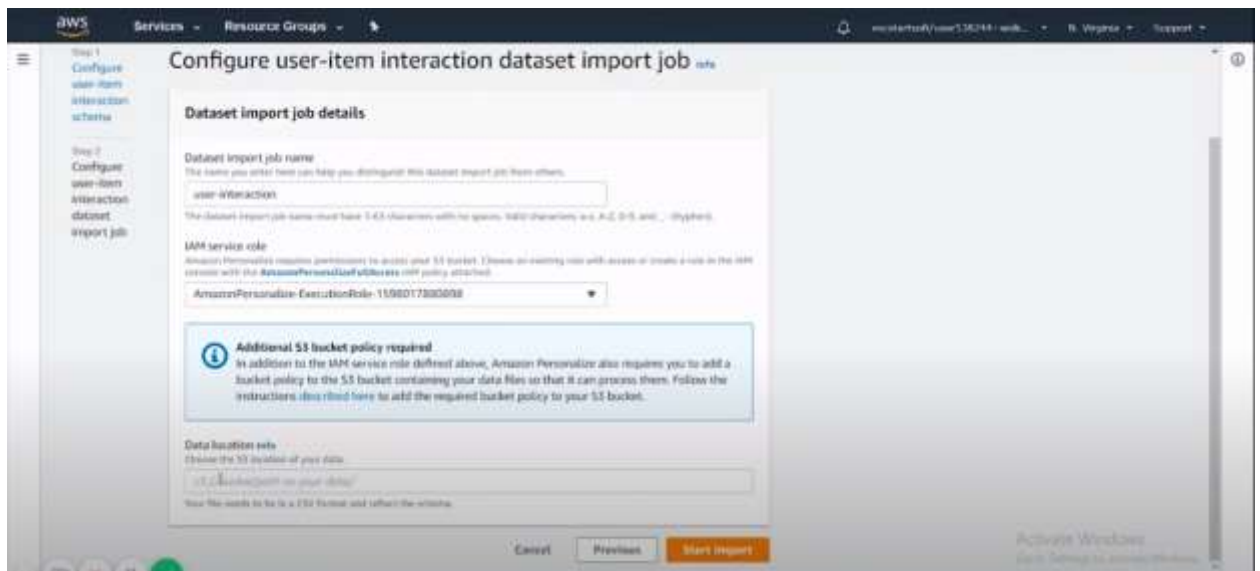
    "namespace": "com.amazonaws.personalize.schema",
    "fields":
    [
      {
        "name": "Ten_cot_1",
        "type": "Kieu_du_lieu_cot_1",
      },
      {
        "name": "Ten_cot_2",
        "type": "Kieu_du_lieu_cot_2",
      }
    ]
    "version": "1.0"
  }
}

```



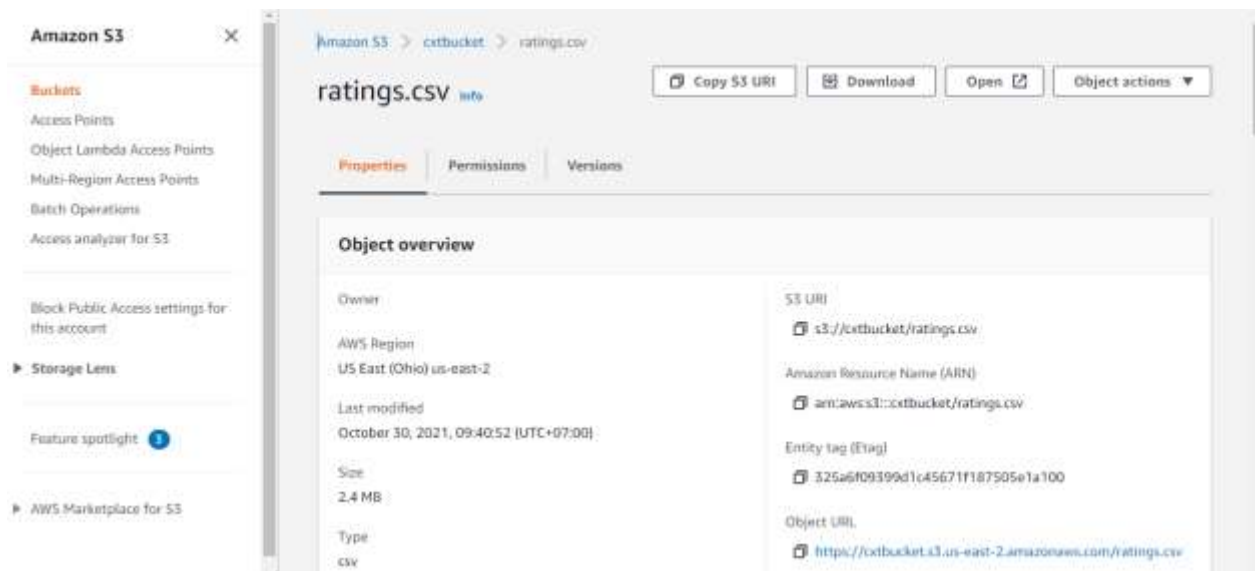
Hình 9 Tạo new schema

Sau khi hoàn thành ta bấm Next -> Nhập tên cho Dataset import job name.

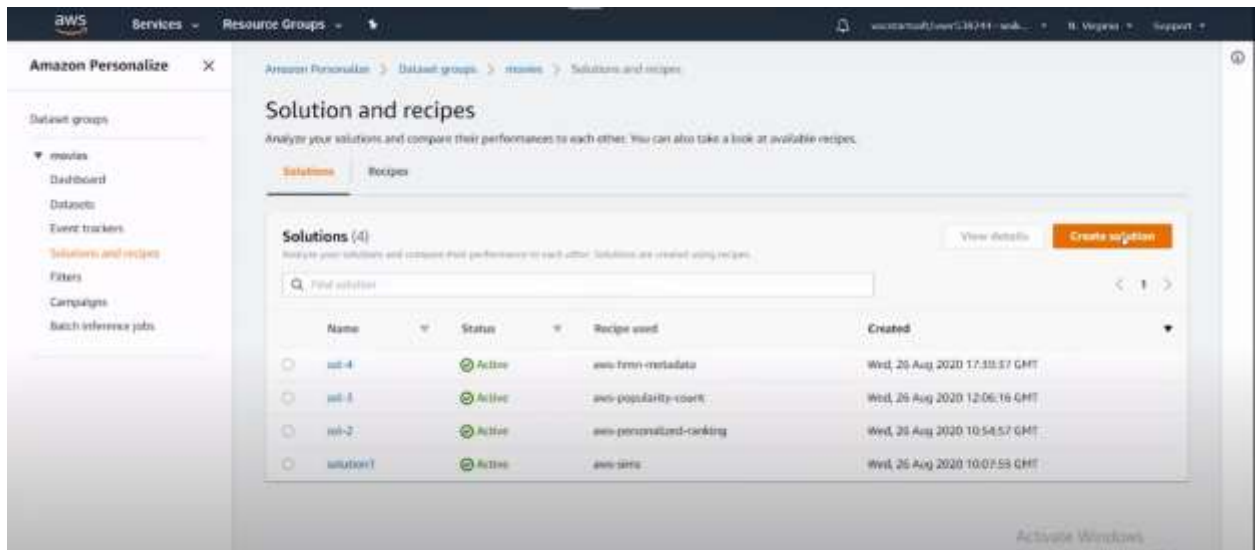


Hình 10 Nhập tên Dataset import job name

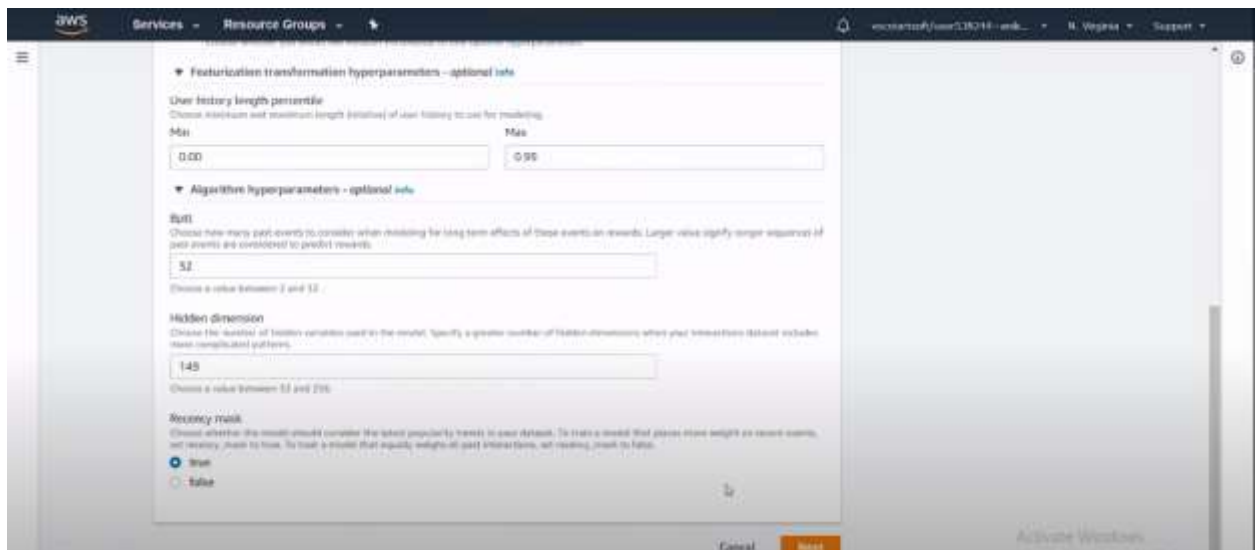
Copy URL S3 -> Paste đường dẫn vào Data location theo thiết kế dataset.



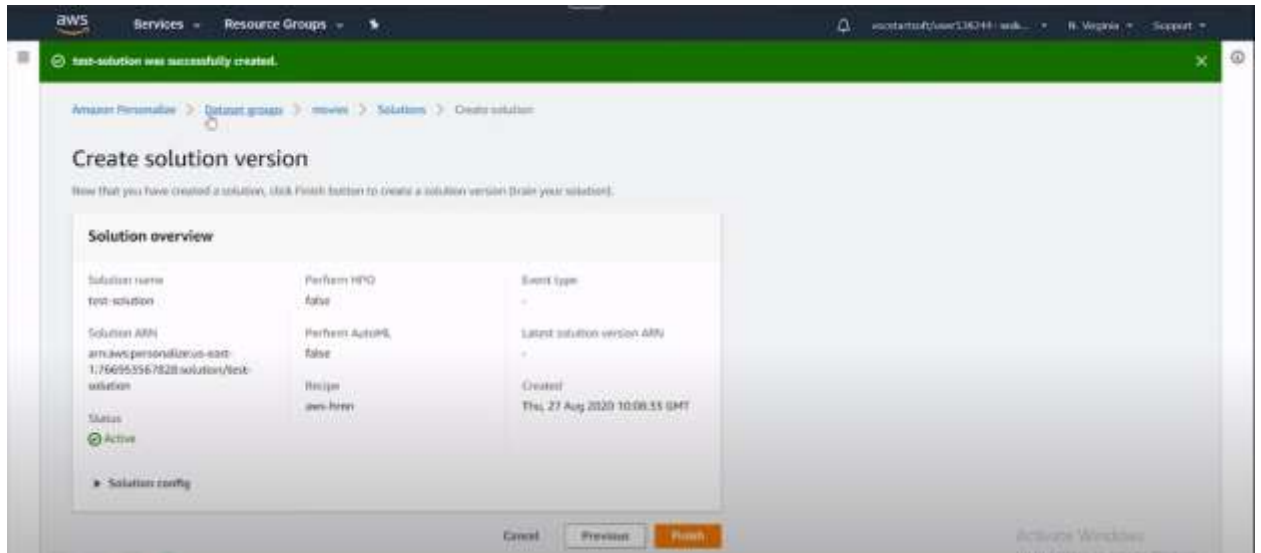
Hình 11 Copy URL S3



Hình 14 Vào Solution and recipes, chọn Create solution

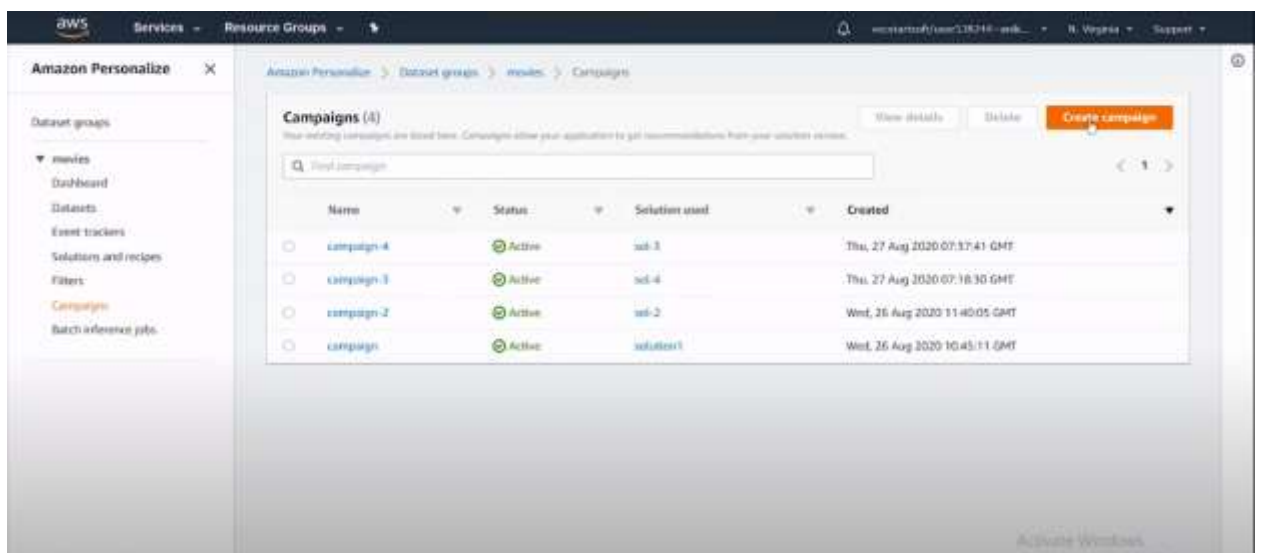


Hình 15 Nhập đầy đủ thông tin và nhấn Next



Hình 16 Kết quả sau khi tạo solution

Tại mục Campaigns -> Bấm vào Create campaign -> Nhập Campaign name, Solution, Solution ID -> Bấm Create campaign.



Hình 17 tại Campaigns chọn Create 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 1-45 characters with no spaces. Valid characters: a-z, A-Z, 0-9, and _ (underscore).

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

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

Wed, 26 Aug 2020 11:01:46 GMT

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

Enter a number from 0-500.

[Cancel](#) [Create campaign](#)

Hình 18 Nhập thông tin và chọn Create campaign

Phần 3: Phần kết

Tóm lại, qua đề tài đồ án mà nhóm chúng em đã làm (tìm hiểu về dịch vụ Recommendation của AWS) đã giúp cho chúng em hiểu được dịch vụ gợi ý (đề xuất) này giúp ta như thế nào cho doanh nghiệp cũng như giúp cho chúng ta trong tương lai. Dịch vụ gợi ý (Recommendation) là hệ thống trí tuệ nhân tạo đem lại những sự gợi ý hợp lý, thích hợp, yêu thích nhất cho người dùng, giúp người dùng tìm kiếm dễ dàng và thích hợp nhất, nó được tích hợp vào nhiều ứng dụng, trang web hay trên cả thị trường thương mại mà từ đó giúp tăng doanh thu cho doanh nghiệp của họ ngày một càng phát triển.

Phần 4: Tài liệu tham khảo

Trang chính AWS

<https://rpubs.com/chidungkt/634300>

<https://viblo.asia/p/gioi-thieu-ve-he-thong-goi-y-recommendation-systems-aWj53kEp56m>

Recommendation System Wikipedia

<https://www.youtube.com/watch?v=Qz9w-DdIJRg>

<https://cloud-ace.vn/dich-vu-va-giai-phap/recommendation-system/>