**Tìm hiểu về dịch vụ Recommendation của AWS**

Trần Minh Sang – 19110278

Cao Xuân Thắng – 19119222

1. **Mô tả**

Đây là demo xậy dựng hệ thống đề xuất bằng Amazon Personalize. Sử dụng data đề xuất là đề xuất cho giáo dục mà phát triển.

Một số thuật toán được sử dụng trong Personalization:

* User-Personalization recipe

Gồm 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

Gồm các thuộc tính sau:

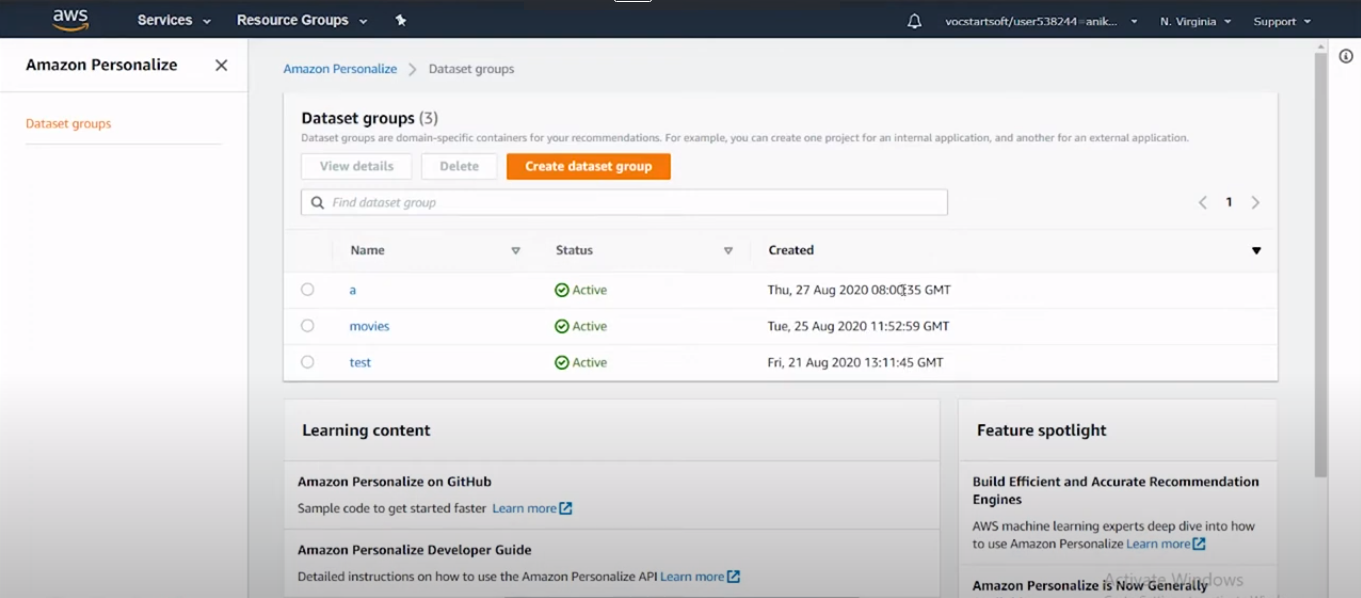
* Name – aws-popularity-count
* Recipe ARN – arn:aws:personalize:::recipe/aws-popularity-count
* Algorithm ARN – arn:aws:personalize:::algorithm/aws-popularity-count
* Feature transformation ARN – arn:aws:personalize:::feature-transformation/sims
* Recipe type – USER\_PERSONALIZATION
* Personalized-Ranking recipe

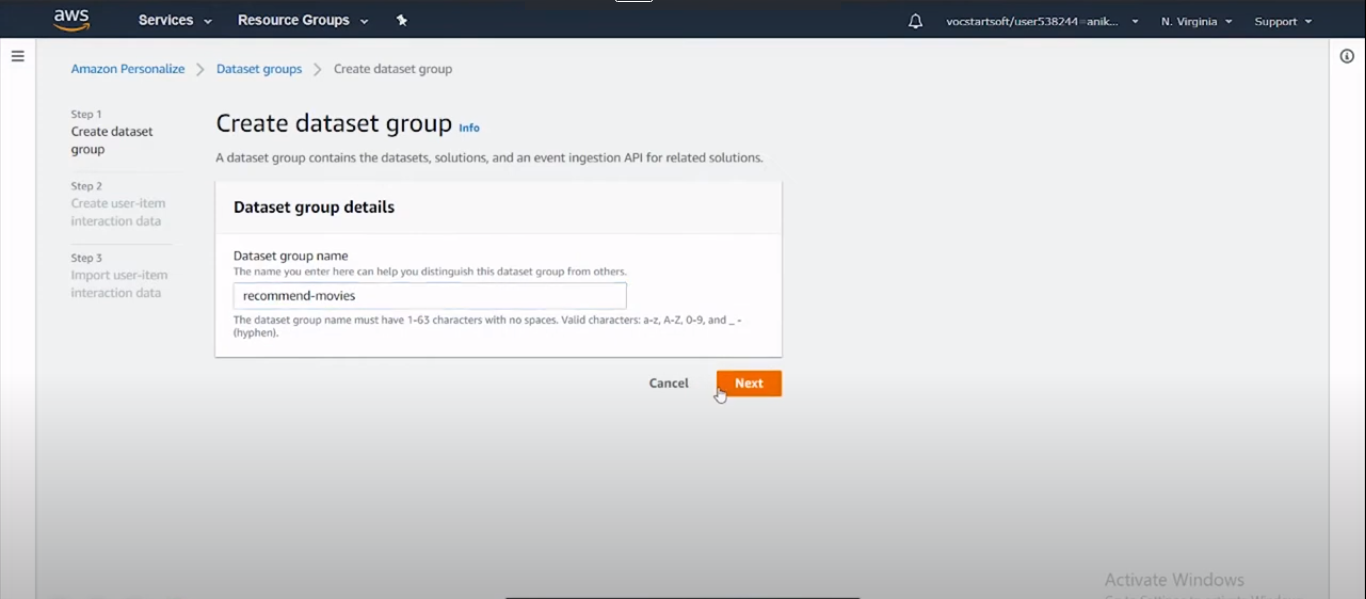
Gồm các thuộc tính sau:

* 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

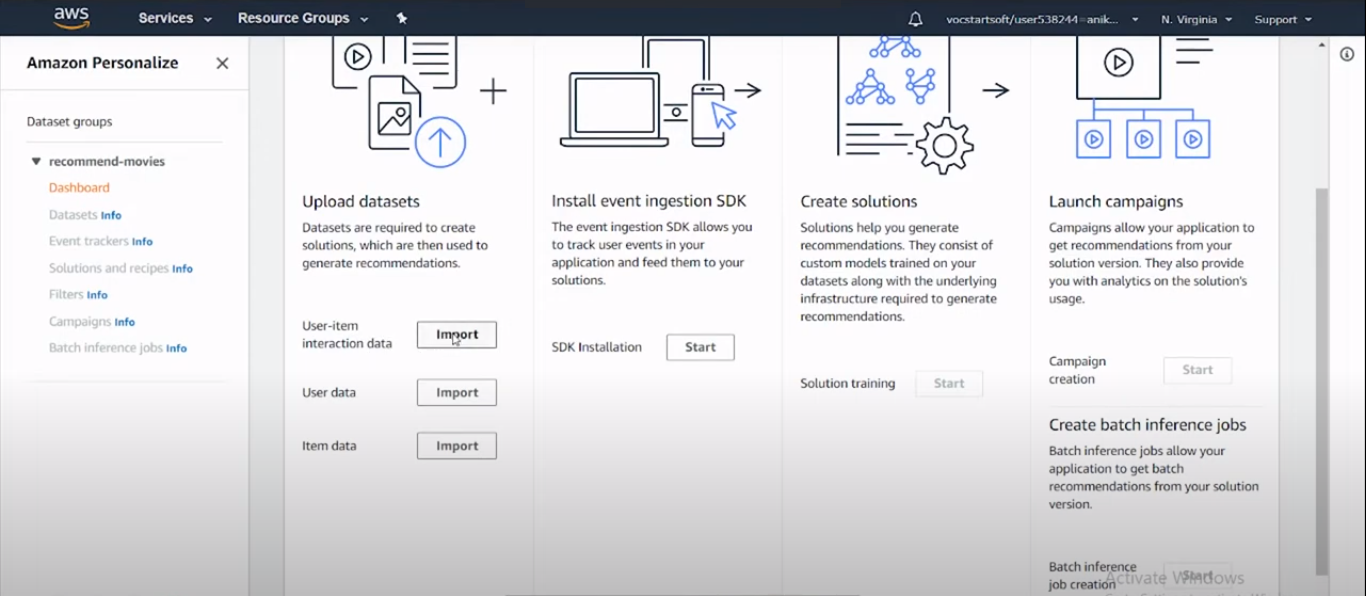
1. **Cách thực hiện**

Đầu tiên chúng ta cần tạo một tài khoản AWS và đăng nhập vào AWS. Sau đó bạn truy cập vào Amazon Personalize của mình -> Bấm vào Create dataset group -> Nhập tên cho Dataset group name -> Bấm Next.





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



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”,

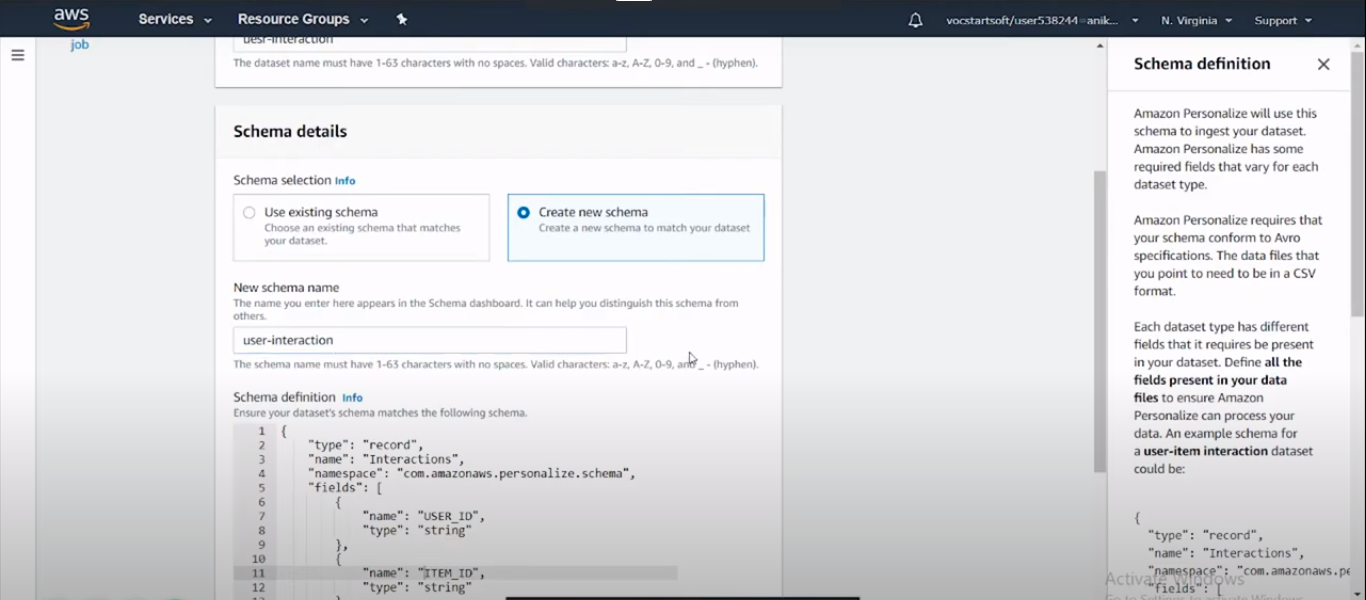
“type”: “Kieu\_du\_lieu\_cot\_2”,

}

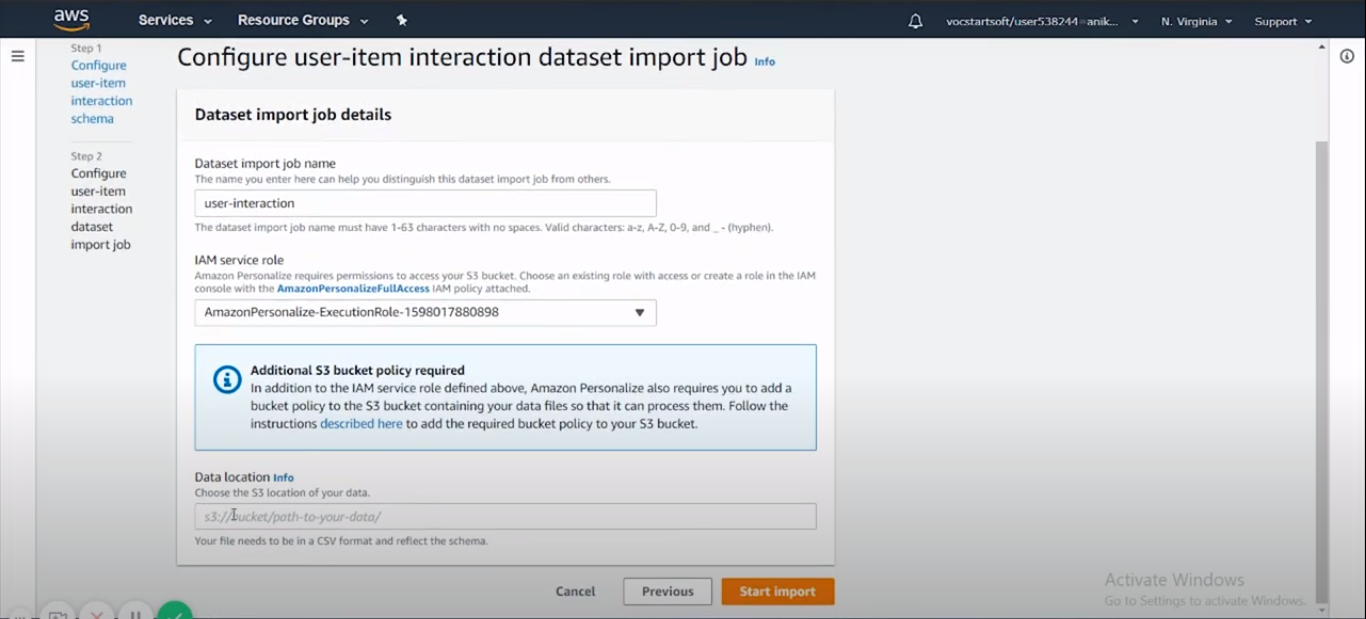
]

“version”: “1.0”

}

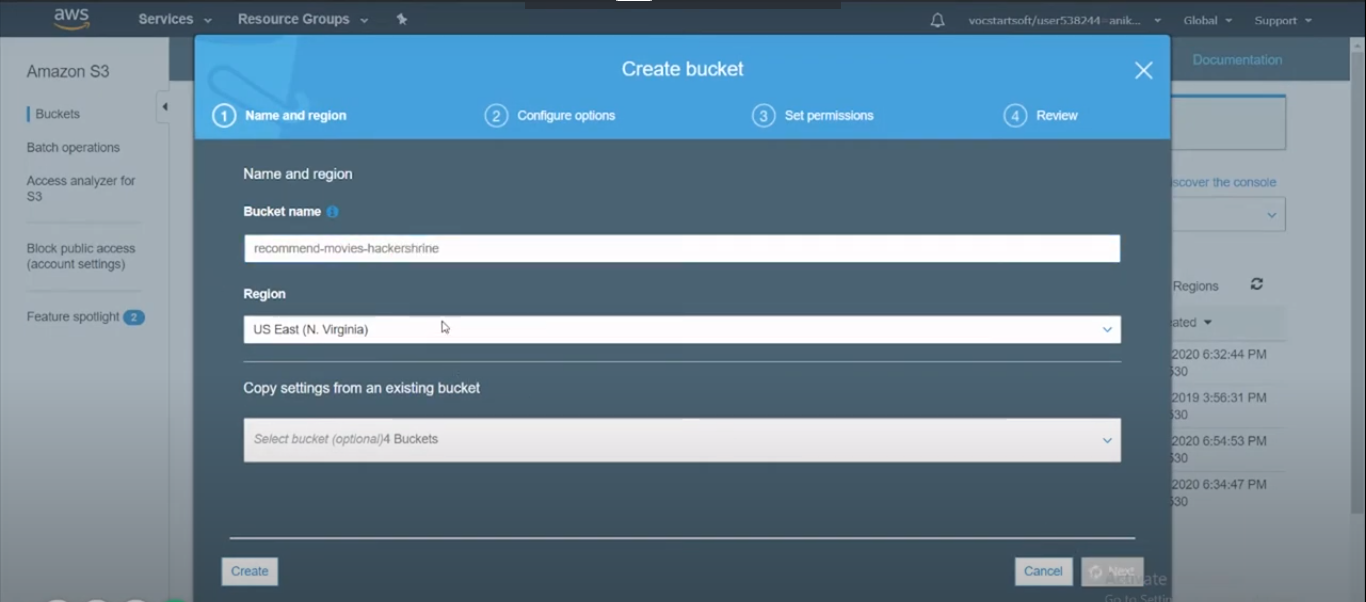


Sau khi hoàn thành ta bấm Next -> Nhập tên cho Dataset inport job name.

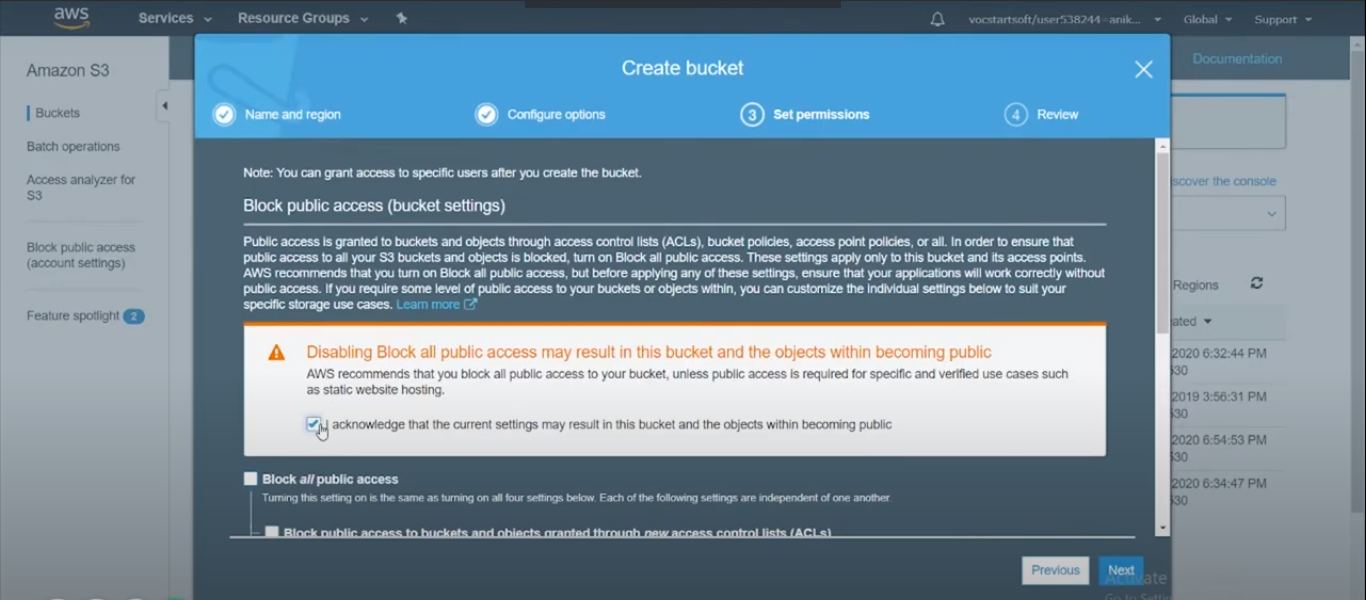


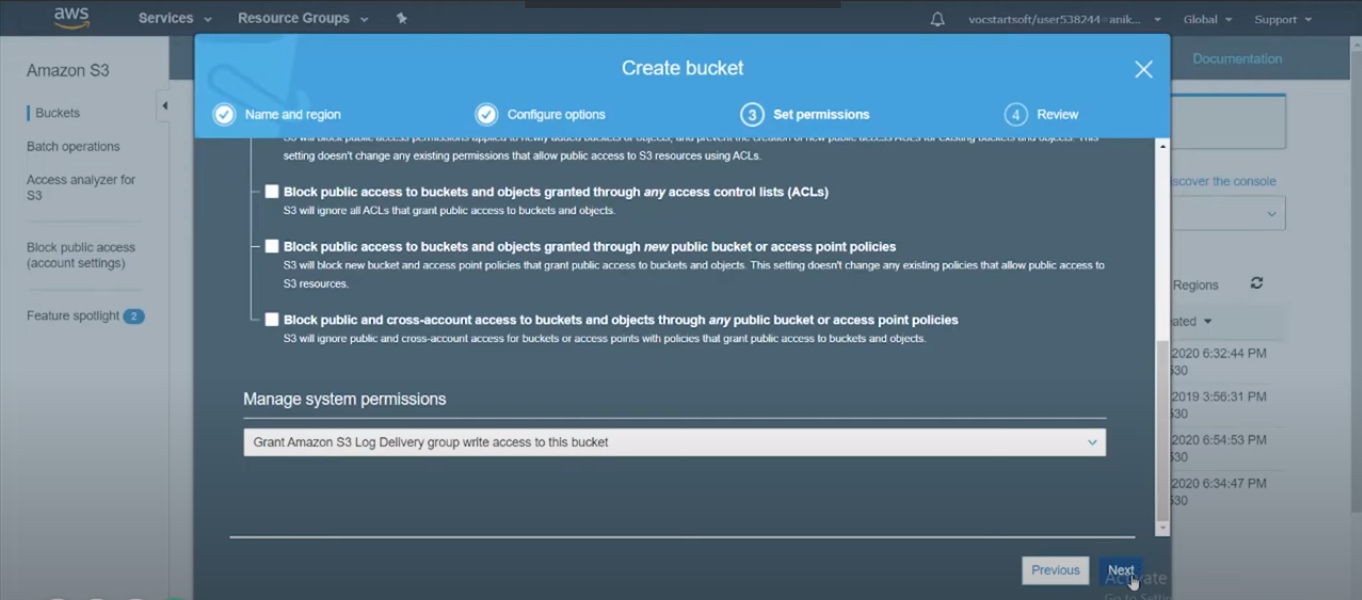
\*Trước khi nhập đường dẫn đến Data location chúng ta phải tạo Buckets tại trang Amazon S3 của mình.

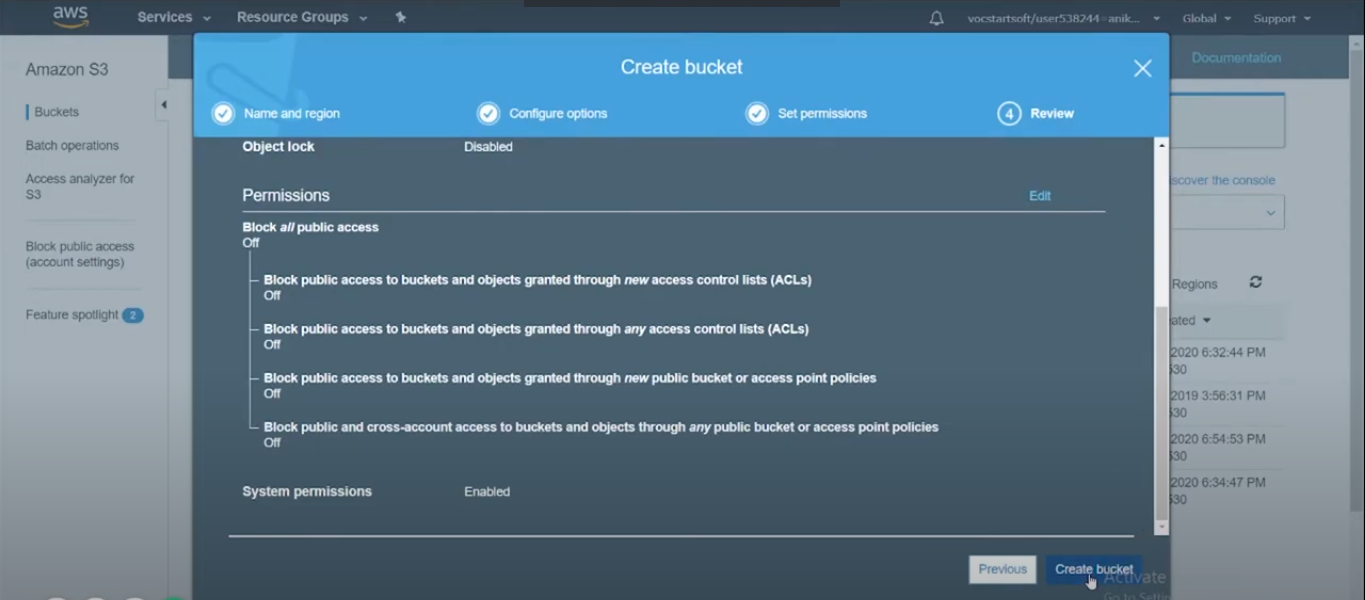
Ta bấm Create bucket -> Nhập Bucket name, Region -> Bấm Next -> Bấm Next -> Bỏ chọn Block all public access -> Phần Manage systerm permissions chọn Grant Amazon S3 Log Delivery group with access to this buckest -> Bấm Next -> Bấm Create bucket.











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",

"Effect": "Allow",

"Action":

[

"s3:GetObject",

"s3:ListBucket"

],

"Resource":

[

"arn:aws:s3:::bucket-name",

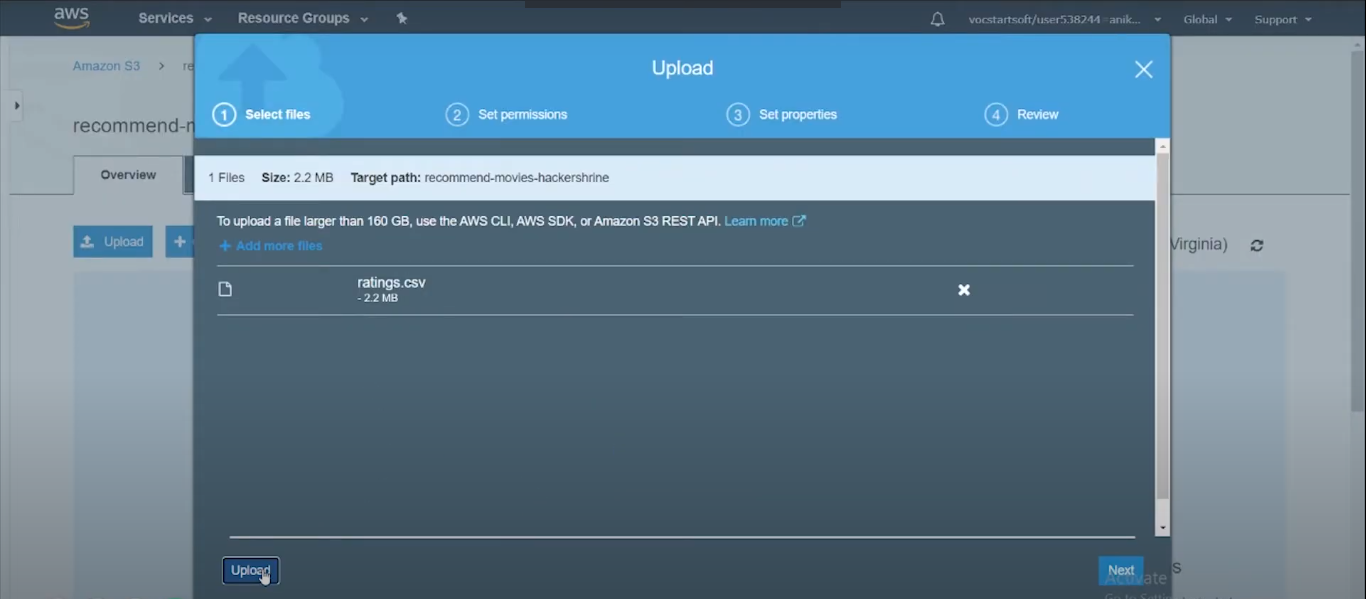
"arn:aws:s3:::bucket-name/\*"

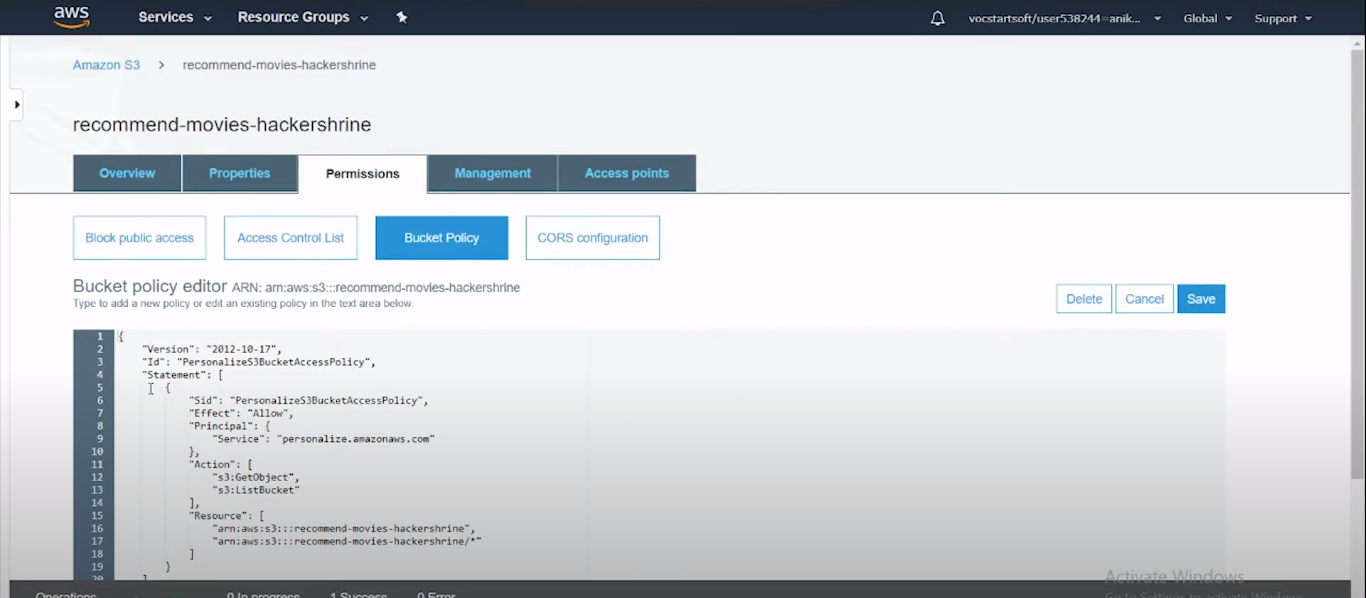
]

}

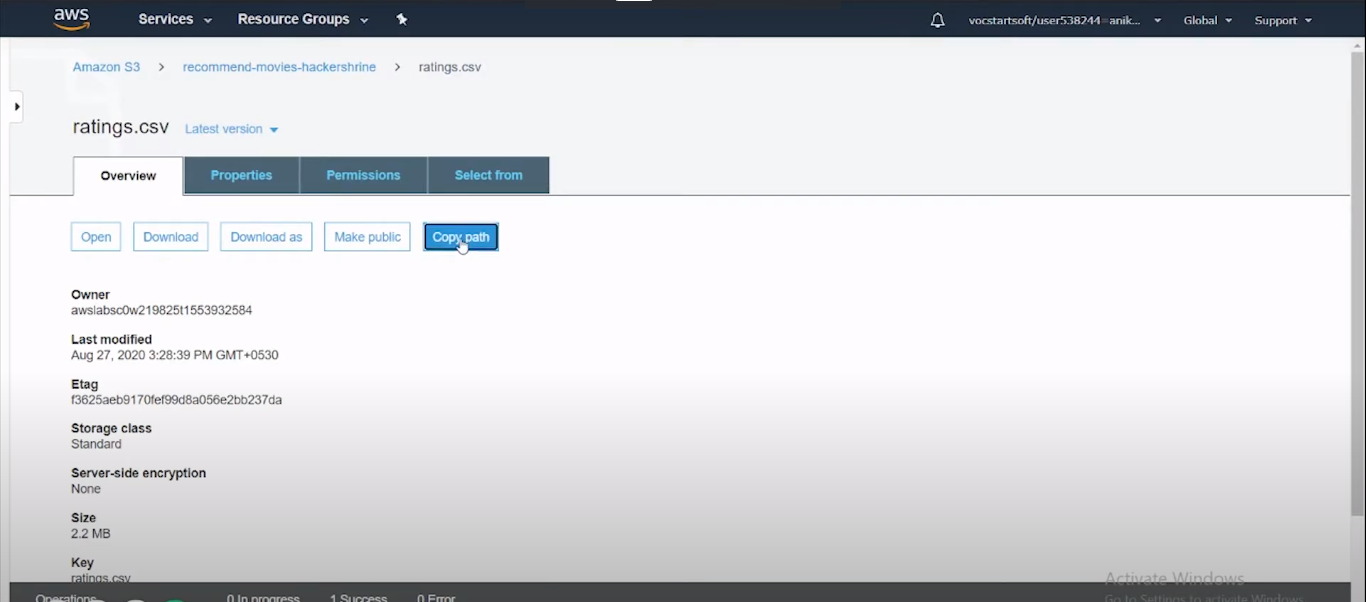
]

}

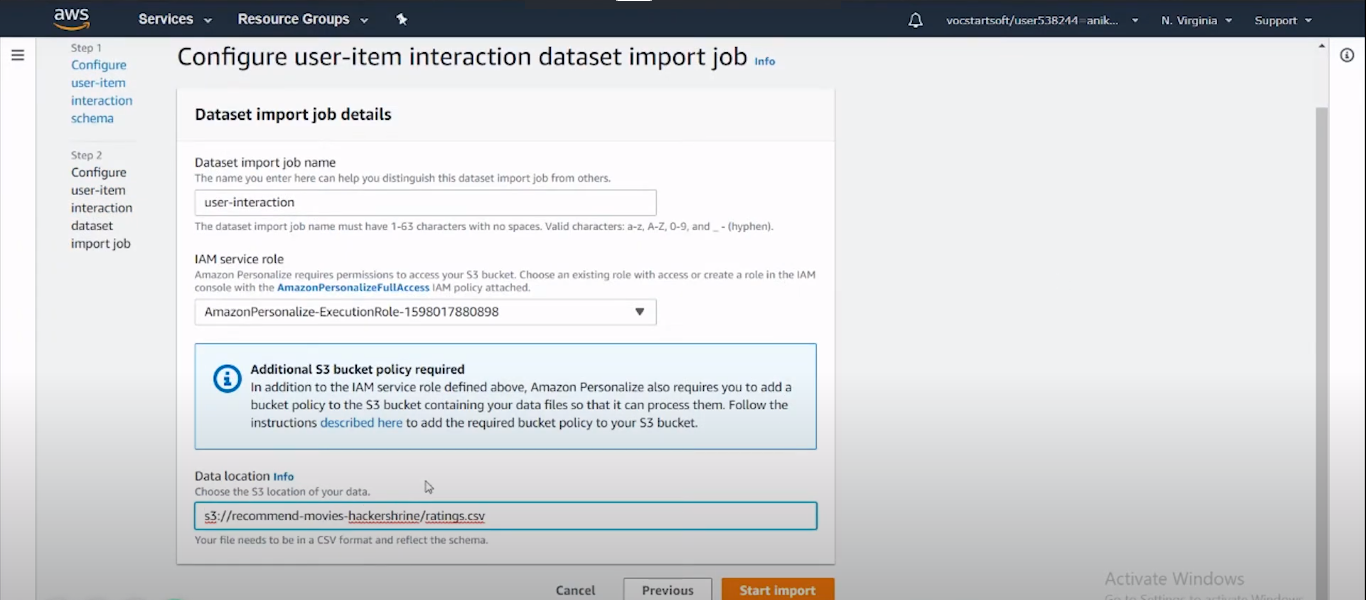




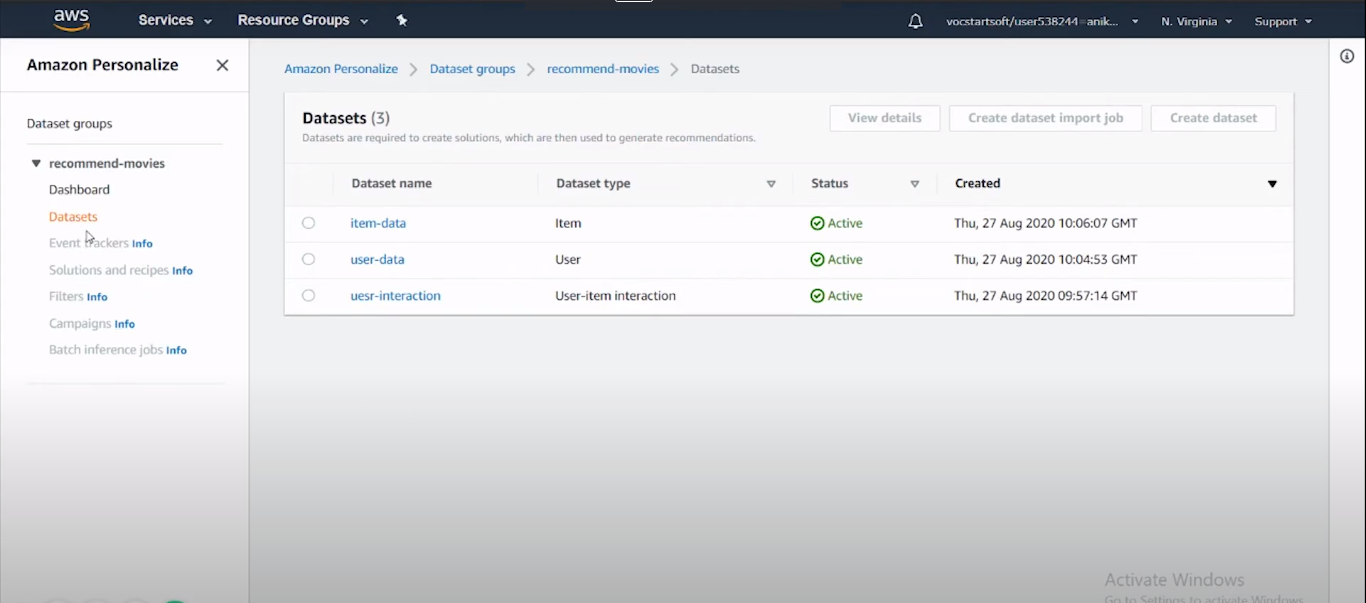
Xong ta vào mục Overview -> Bấm vào file vừa tải lên -> Bấm vào Copy path.



Ta đã tạo xong Bucket. Quay lại trang ở bước 4 -> Paste đường dẫn vào Data location -> Bấm Start inport



Quay lại bước 3-4-6-7-8 (không cần tạo thêm Bucket nữa) với hai phần còn lại là User data và Item data. Lưu ý ở bước 4 ta phải thiết kế Schema definition phù hợp với file mình load lên từ máy.

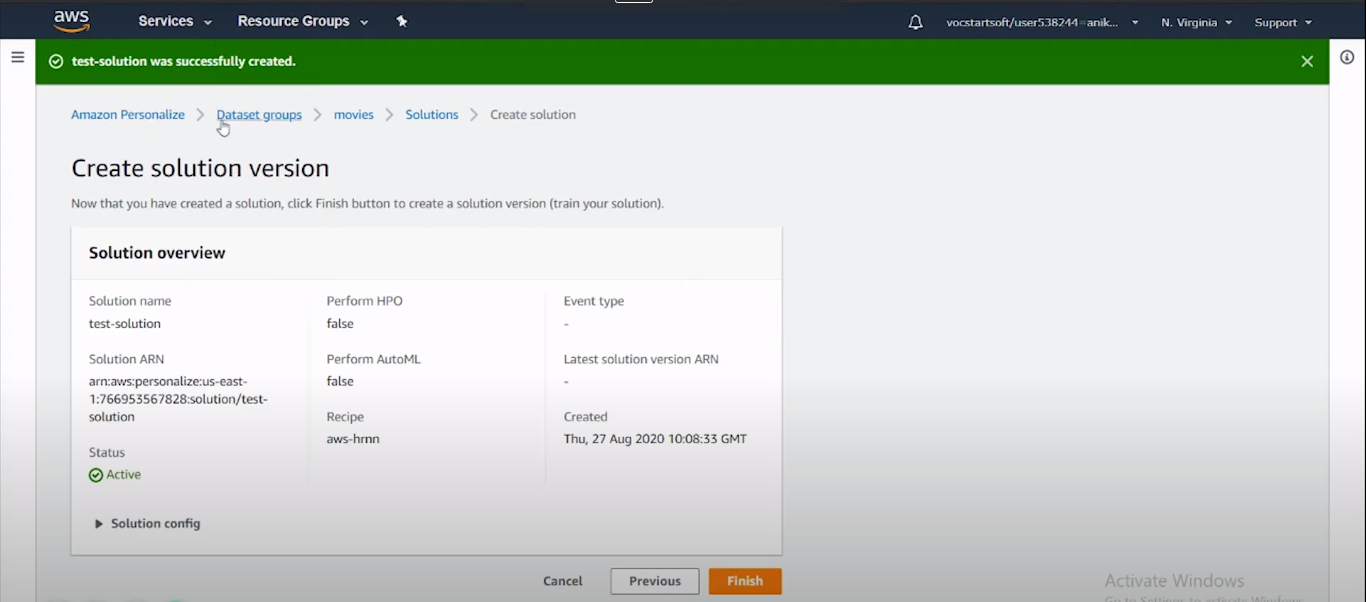


Đây là kết quả

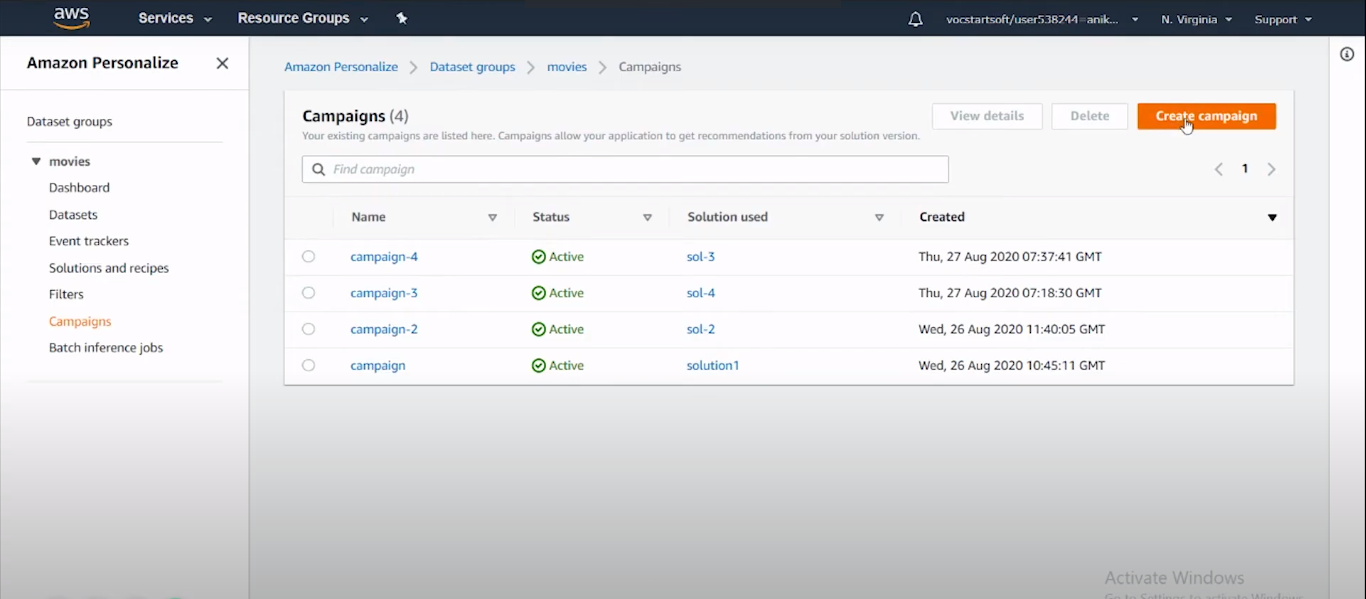
Tại mục Solutions and recipes -> Bấm vào Create solution -> Nhập Solution name -> Chọn thuật toán bạn muốn tại Recipe -> Thay đổi cái thiết lập theo mong muốn tại Advanced configuration -> Bấm Next -> Bấm Finish.

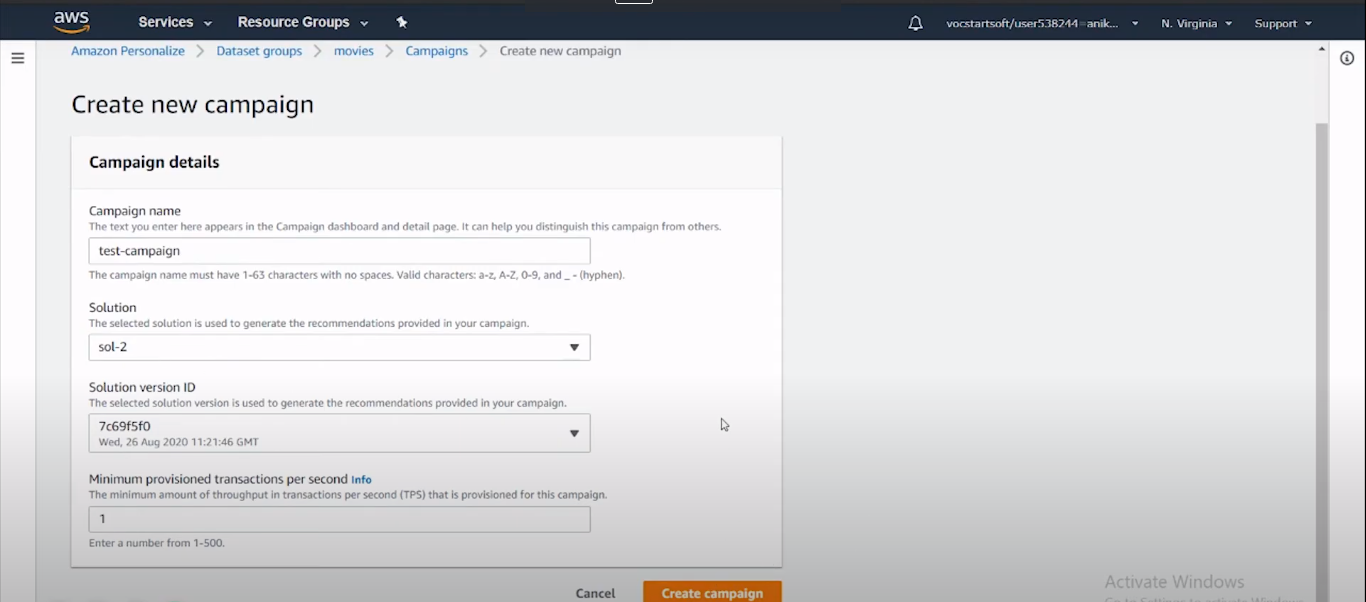






Tại mục Campaigns -> Bấm vào Create campaign -> Nhập Campaign name, Solution, Solution ID -> Bấm Create campaign.





VD:

- “Recommend\_movie” đề suất bộ phim phù hợp với bản thân mình nhất dựa theo các yếu tố như: thể loại, thời lượng, điểm đánh giá, số lượng người coi.

- Thết kế Dataset group như sau:

User-item interaction data: ratings.csv.

User data: tags.csv.

Item data: movies.csv, links.csv.