

The background features a light gray map of Thailand on the right side, composed of numerous small, semi-transparent icons representing various digital and social media concepts like hearts, search magnifying glasses, musical notes, video cameras, thumbs up, and social media logos. A thin green horizontal line is positioned above the title, and a thin orange horizontal line is positioned below it.

A Scalable Cloud Architecture for Continuous Fine-Tuning and Deployment of LLMs on Big Data Streams

Tares Supadol
6709300120

Background

2

ในยุคของการค้าขายออนไลน์ (E-commerce) ปริมาณข้อมูลจากผู้บริโภคเติบโตอย่างรวดเร็ว โดยเฉพาะข้อมูลในรูปแบบ **ข้อความรีวิวสินค้า (customer reviews)** ที่ผู้ใช้งานทิ้งไว้หลังการซื้อ ซึ่งเป็นข้อมูลที่มีคุณค่าอย่างมากในการวิเคราะห์ความพึงพอใจ ความต้องการ และประสบการณ์ที่ลูกค้ามีต่อสินค้า

Amazon เป็นหนึ่งในแพลตฟอร์ม E-commerce ที่มีรีวิวสินค้าจำนวนมากมหาศาล โดยในแต่ละวันมีข้อมูลใหม่ ๆ ไหลเข้ามาอย่างต่อเนื่อง การจัดการข้อมูลประเภทนี้ จำเป็นต้องมีระบบที่สามารถ **วิเคราะห์แบบอัตโนมัติ, ปรับปรุงความแม่นยำของโมเดลต่อเนื่อง และตอบสนองในระดับ real-time** เพื่อให้ธุรกิจสามารถปรับปรุงสินค้า บริการ และการตลาดได้อย่างทันท่วงที

Background

Compared to Previous Versions

Year	#Review	#User	#Item	#R_Token	#M_Token	#Domain	Timespan
2013	34.69M	6.64M	2.44M	5.91B	–	28	Jun'96 - Mar'13
2014	82.83M	21.13M	9.86M	9.16B	4.14B	24	May'96 - Jul'14
2018	233.10M	43.53M	15.17M	15.73B	7.99B	29	May'96 - Oct'18
2023	571.54M	54.51M	48.19M	30.14B	30.78B	33	May'96 - Sep'23

Grouped by Category

See also

Check Pure ID files ("ratings only" files) and corresponding data splitting strategies in [Common Data Processing](#) section.

Category	#User	#Item	#Rating	#R_Token	#M_Token	Download
All_Beauty	632.0K	112.6K	701.5K	31.6M	74.1M	review, meta
Amazon_Fashion	2.0M	825.9K	2.5M	94.9M	510.5M	review, meta
Appliances	1.8M	94.3K	2.1M	92.8M	95.3M	review, meta
Arts_Crafts_and_Sewing	4.6M	801.3K	9.0M	350.0M	695.4M	review, meta
Automotive	8.0M	2.0M	20.0M	824.9M	1.7B	review, meta
Baby_Products	3.4M	217.7K	6.0M	323.3M	218.6M	review, meta
Beauty_and_Personal_Care	11.3M	1.0M	23.9M	1.1B	913.7M	review, meta
Books	10.3M	4.4M	29.5M	2.9B	3.7B	review, meta
CDs_and_Vinyl	1.8M	701.7K	4.8M	514.8M	287.5M	review, meta
Cell_Phones_and_Accessories	11.6M	1.3M	20.8M	935.4M	1.3B	review, meta
Clothing_Shoes_and_Jewelry	22.6M	7.2M	66.0M	2.6B	5.9B	review, meta
Digital_Music	101.0K	70.5K	130.4K	11.4M	22.3M	review, meta
Electronics	18.3M	1.6M	43.9M	2.7B	1.7B	review, meta
Gift_Cards	132.7K	1.1K	152.4K	3.6M	630.0K	review, meta

ทำไมถึงอยากทำ / ได้อะไร

- ปัจจุบันองค์กรมีข้อมูลข้อความขนาดใหญ่ที่ต้องการวิเคราะห์
- LLMs มีศักยภาพสูงในการประมวลผลภาษามนุษย์
- งานวิจัยนี้จะสร้างระบบที่เรียนรู้จากข้อมูลใหม่และนำไปใช้งานได้จริง

1. รับข้อมูลรีวิวลูกค้าจาก Amazon Customer Reviews Dataset อย่างต่อเนื่อง
2. ประมวลผลและเตรียมข้อมูลให้อยู่ในรูปแบบที่พร้อมฝึกโมเดล (LLM)
3. ทำการ ฝึกโมเดลภาษาขนาดใหญ่ (LLM) เช่น DistilBERT หรือ T5 เพื่อวิเคราะห์ sentiment ของผู้บริโภค
4. ทำการ ฝึกซ้ำ (Fine-tune) โมเดลอัตโนมัติ เมื่อมีข้อมูลใหม่เข้ามา
5. ให้บริการผ่าน API ที่สามารถทำนาย sentiment แบบ real-time
6. แสดงผลการทำงานและแนวโน้มของความรู้สึกของลูกค้าใน Dashboard

วัตถุประสงค์

6

ตารางสรุป KPI ของระบบ เช่น:

KPI	เป้าหมาย
-----	----------

ความแม่นยำ (Accuracy)	>90%
-----------------------	------

Latency ของ API	<500ms
-----------------	--------

เวลา Retrain โมเดล	ทุกสัปดาห์
--------------------	------------

Model

7

โมเดลที่ใช้	ผลลัพธ์ที่คาดว่าจะได้จากโมเดล	ใช้ทำอะไร
DistilBERT / BERT	ทำนายว่า ข้อความนั้นเป็นบวก / ลบ / กลาง (sentiment)	วิเคราะห์รีวิวสินค้า / ติดตามความคิดเห็นลูกค้า
RoBERTa	ทำนายผลด้วยความแม่นยำสูงขึ้น (คล้าย BERT)	ใช้ในงานที่ต้องการประสิทธิภาพสูง เช่น คัดกรอง feedback ที่สำคัญ
T5 / ThaiT5 (Text-to-Text Transfer Transformer)	สรุปข้อความ / แปลภาษา / ตอบคำถาม จาก input ที่กำหนด	สรุปรีวิวให้สั้นลงเพื่ออ่านง่าย / ตอบคำถามจากข้อความยาว
WangchanBERTa / Thai2Fit	เข้าใจข้อความภาษาไทยแบบละเอียด เช่น ภาษาพูด	ใช้กับข้อมูลภาษาไทย เช่น รีวิว Shopee / Facebook / Pantip

Exam from the model

1. Sentiment Dashboard แบบ Real-time

แสดงว่าผู้ใช้ส่วนใหญ่ “พอใจหรือไม่พอใจ” กับสินค้าหรือบริการ

2. ระบบเตือนอัตโนมัติ

หากมีรีวิวเชิงลบเพิ่มขึ้นในหมวดสินค้าใด → ระบบสามารถแจ้งเตือนทีมดูแล

3. การวัดประสิทธิภาพการสื่อสาร / ผลลัพธ์

เปรียบเทียบ Sentiment ก่อนและหลังเปลี่ยนผลลัพธ์ หรือแคมเปญโฆษณา

4. สรุปความคิดเห็นของลูกค้า

เช่น สรุปข้อความ 500 รีวิว ให้เหลือแค่ 3 ประโยคเข้าใจง่าย (ถ้าใช้ T5)

Evaluation Method

9

- Accuracy / F1-score ของโมเดลในการทำนาย sentiment
- Latency ของการทำนายผ่าน API
- เวลาในการ retrain โมเดลแต่ละครั้ง
- ความถูกต้องของการเรียนรู้แบบต่อเนื่อง (Continuous Learning Accuracy Drift)
- ความพึงพอใจของผู้ใช้ (User Feedback บน Dashboard)

- ชุดข้อมูล: **Amazon Customer Reviews Dataset**
(<https://nijianmo.github.io/amazon/index.html>)
- ข้อมูลรีวิวกจากหมวด **Electronics** หรือ **Books**
- ฟิลด์: review Text, rating, summary, review Time, Overall, ProductTitle
- ขนาดใหญ่เพียงพอสำหรับงานวิจัย
- สามารถแบ่ง batch สำหรับการ retraining ได้

5 Vs of Big Data

11

หัวข้อ	รายละเอียด
Volume ปริมาณ	ข้อมูลมีขนาดใหญ่ระดับ หลักแสนถึงล้านรายการ เช่น รีวิวลูกค้าใน Amazon ซึ่งมีมากพอให้ฝึกโมเดล LLM และทดสอบการทำงานของระบบที่ขยายขนาดได้ (20,994,353 reviews)
Velocity ความเร็ว	ข้อมูลรีวิวถูกเพิ่มเข้ามา อย่างต่อเนื่อง (stream) เช่น รีวิวใหม่ที่เกิดขึ้นทุกวัน → ใช้สร้างระบบ Continuous Fine-Tuning ที่เรียนรู้จากข้อมูลใหม่
Variety ความหลากหลาย	ข้อมูลมีหลายรูปแบบ: ข้อความ (reviewText), คะแนน (rating), วันที่ (reviewTime) ฯลฯ
Veracity ความเชื่อถือได้	ข้อมูลมี noise เช่น รีวิวสั้นเกินไป / ข้อมูลไม่ครบ ต้องมีขั้นตอน Data Cleaning เพื่อให้ได้ข้อมูลที่เหมาะกับการฝึกโมเดล
Value คุณค่า	การใช้ LLM วิเคราะห์ Sentiment หรือสรุปความคิดเห็นของลูกค้า ช่วยสร้าง Insight และช่วยให้ธุรกิจตัดสินใจได้ดีขึ้น → สร้างคุณค่าเชิงธุรกิจ จาก Big Data

5 Vs of Big Data

12

raw/

Copy S3 URI

Objects

Properties

Objects (5)



Copy S3 URI

Copy URL

Download

Open

Delete

Actions

Create folder

Upload

Objects are the fundamental entities stored in Amazon S3. You can use [Amazon S3 inventory](#) to get a list of all objects in your bucket. For others to access your objects, you'll need to explicitly grant them permissions. [Learn more](#)

Find objects by prefix

Show versions

< 1 > ⚙

<input type="checkbox"/>	Name	Type	Last modified	Size	Storage class
<input type="checkbox"/>	Books.json.gz	gz	May 17, 2025, 17:35:03 (UTC+07:00)	10.9 GB	Standard
<input type="checkbox"/>	Electronics.json	json	May 17, 2025, 07:15:14 (UTC+07:00)	10.7 GB	Standard
<input type="checkbox"/>	meta_Books.json	json	May 17, 2025, 11:10:32 (UTC+07:00)	3.8 GB	Standard
<input type="checkbox"/>	meta_Electronics.json	json	May 17, 2025, 11:13:04 (UTC+07:00)	10.2 GB	Standard
<input type="checkbox"/>	small.json	json	May 30, 2025, 23:28:42 (UTC+07:00)	6.1 MB	Standard

ชุดข้อมูลและรายละเอียดของชุดข้อมูลที่ใช้

13

Complete review data

Please only download these (large!) files if you really need them. We recommend using the smaller datasets (i.e. k-core and CSV files) as shown in the [next section](#).

raw review data (34gb) - all 233.1 million reviews

ratings only (6.7gb) - same as above, in csv form without reviews or metadata

5-core (14.3gb) - subset of the data in which all users and items have at least 5 reviews (75.26 million reviews)

meta data (12gb) - meta data for all products

- We also provide a [colab notebook](#) that helps you parse and clean the data.

Per-category data - the review and product metadata for each category:

Amazon Fashion	reviews (883,636 reviews)	metadata (186,637 products)
All Beauty	reviews (371,345 reviews)	metadata (32,992 products)
Appliances	reviews (602,777 reviews)	metadata (30,459 products)
Arts Crafts and Sewing	reviews (2,875,917 reviews)	metadata (303,426 products)
Automotive	reviews (7,990,166 reviews)	metadata (932,019 products)
Books	reviews (51,311,621 reviews)	metadata (2,935,525 products)
CDs and Vinyl	reviews (4,543,369 reviews)	metadata (544,442 products)
Cell Phones and Accessories	reviews (10,063,255 reviews)	metadata (590,269 products)
Clothing Shoes and Jewelry	reviews (32,292,099 reviews)	metadata (2,685,059 products)
Digital Music	reviews (1,584,082 reviews)	metadata (465,392 products)
Electronics	reviews (20,994,353 reviews)	metadata (786,868 products)
Gift Cards	reviews (147,194 reviews)	metadata (1,548 products)
Grocery and Gourmet Food	reviews (5,074,160 reviews)	metadata (287,209 products)
Home and Kitchen	reviews (21,928,568 reviews)	metadata (1,301,225 products)
Industrial and Scientific	reviews (1,758,333 reviews)	metadata (167,524 products)
Kindle Store	reviews (5,722,988 reviews)	metadata (493,859 products)
Luxury Beauty	reviews (574,628 reviews)	metadata (12,308 products)
Magazine Subscriptions	reviews (89,689 reviews)	metadata (3,493 products)
Movies and TV	reviews (8,765,568 reviews)	metadata (203,970 products)

ชุดข้อมูลและรายละเอียดของชุดข้อมูลที่ใช้

14

Name	Size	Date modified	Name	Size	Date modified	Type
Books.json		3/5/2568 16:34	Books.json	36,766,034 KB	17/1/2568 11:55	JSON Source File
Electronics.json		3/5/2568 16:34	Electronics.json	11,238,741 KB	17/1/2568 12:09	JSON Source File
meta_Books.json		3/5/2568 16:34	meta_Books.json	3,958,320 KB	17/1/2568 15:53	JSON Source File
meta_Electronics.json		3/5/2568 16:30	meta_Electronics.json	10,737,922 KB	17/1/2568 15:58	JSON Source File
Books.json.gz	11,422,907 KB	3/5/2568 16:34				
Books_5.json.gz	6,924,962 KB	3/5/2568 18:45				
Electronics.json.gz	3,203,793 KB	3/5/2568 16:08				
meta_Books.json.gz	1,190,413 KB	3/5/2568 16:13				
meta_Electronics.json.gz	1,174,903 KB	3/5/2568 16:11				

เอกสาร WinRAR

เอกสาร WinRAR

ชุดข้อมูลและรายละเอียดของชุดข้อมูลที่ใช้

15

ก่อนทำ ETL

raw/

Objects Properties

Objects (5)

Objects are the fundamental entities stored in Amazon S3. You can use [Amazon S3 inventory](#) to get a list of all objects in your bucket. For others to access your objects, you'll need to explicitly grant them permissions.

Find objects by prefix

Show versions

<input type="checkbox"/>	Name	Type	Last modified	Size	Storage class
<input type="checkbox"/>	Books.json.gz	gz	May 17, 2025, 17:35:03 (UTC+07:00)	10.9 GB	Standard
<input type="checkbox"/>	Electronics.json	json	May 17, 2025, 07:15:14 (UTC+07:00)	10.7 GB	Standard
<input type="checkbox"/>	meta_Books.json	json	May 17, 2025, 11:10:32 (UTC+07:00)	3.8 GB	Standard
<input type="checkbox"/>	meta_Electronics.json	json	May 17, 2025, 11:13:04 (UTC+07:00)	10.2 GB	Standard
<input type="checkbox"/>	small.json	json	May 30, 2025, 23:28:42 (UTC+07:00)	6.1 MB	Standard

Upload

Add the files and folders you want to upload to S3. To upload a file larger than 160GB, use the AWS CLI, AWS SDKs or Amazon S3 REST API. [Learn more](#)

Drag and drop files and folders you want to upload here, or choose [Add files](#) or [Add folder](#).

Files and folders (1 total, 6.1 MB)

All files and folders in this table will be uploaded.

Find by name

<input type="checkbox"/>	Name	Folder	Type	Size
<input type="checkbox"/>	small.json	-	application/json	6.1 MB

Destination

Destination

s3://s3-0120-68/etl/

Destination details

Bucket settings that impact new objects stored in the specified destination.

Permissions

Grant permissions and access to other AWS accounts.

Properties

Specify storage class, encryption settings, tags, and more.

หลังทำ ETL

electronic_db

Last updated (UTC)
May 30, 2025 at 16:59:11

Database properties

Name	Description	Location	Created on (UTC)
electronic_db	-	-	May 30, 2025 at 16:37:08

Tables (1)

View and manage all available tables.

Last updated (UTC)
May 30, 2025 at 16:59:14

Delete

Add tables using crawler

Add table

Filter tables

<input type="checkbox"/>	Name	Database	Location	Classification	Deprecated	View data	Data quality	Column statistics
<input type="checkbox"/>	etl	electronic_db	s3://s3-0120-68/etl/	CSV	-	Table data	View data quality	View statistics

รายละเอียดชุดข้อมูล

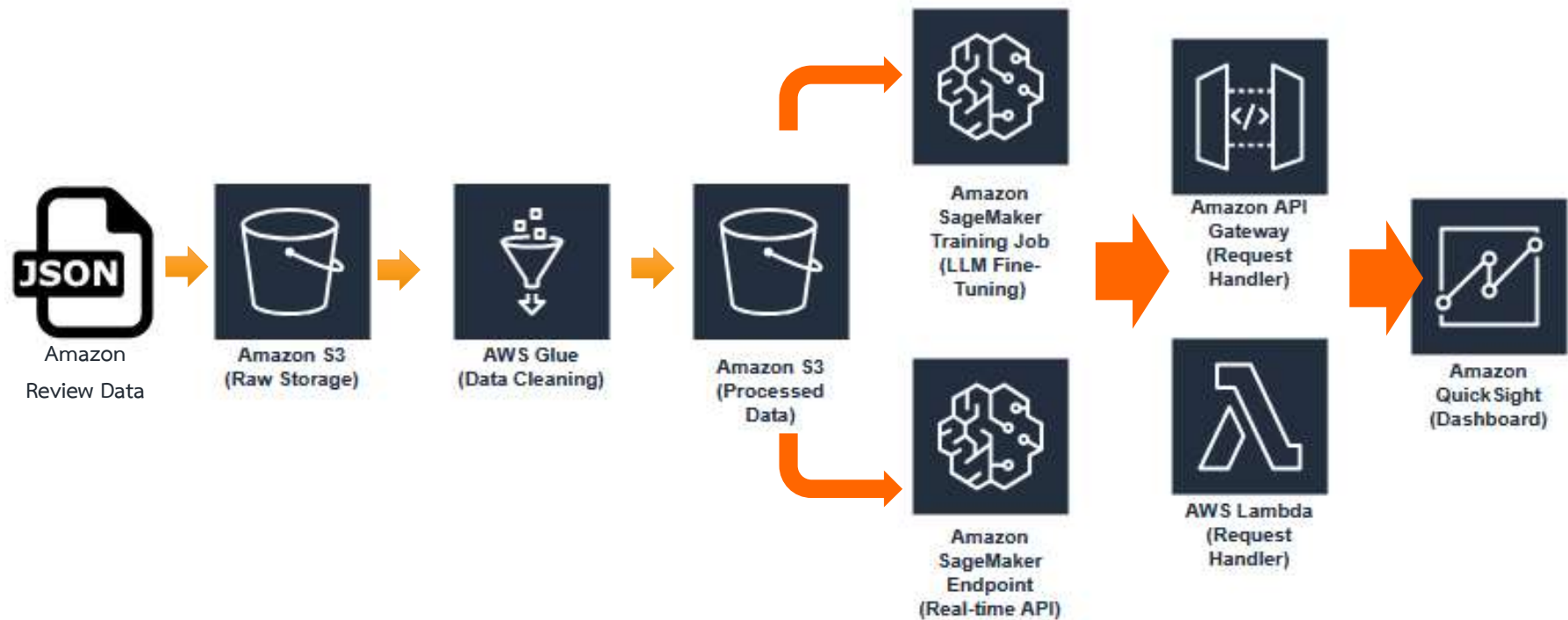
- reviewerID - ID of the reviewer, e.g. A2SUAM1J3GNN3B
- asin - ID of the product, e.g. 0000013714
- reviewerName - name of the reviewer
- vote - helpful votes of the review
- style - a dictionary of the product metadata, e.g., "Format" is "Hardcover"
- reviewText - text of the review
- overall - rating of the product
- summary - summary of the review
- unixReviewTime - time of the review (unix time)
- reviewTime - time of the review (raw)
- image - images that users post after they have received the product

ตัวอย่างชุดข้อมูล

```
{
  "image": ["https://images-na.ssl-images-amazon.com/images/I/71eG75FTJJL._SY88.jpg"],
  "overall": 5.0,
  "vote": "2",
  "verified": True,
  "reviewTime": "01 1, 2018",
  "reviewerID": "AUI6WTTT0QZYS",
  "asin": "5120053084",
  "style": {
    "Size": "Large",
    "Color": "Charcoal"
  },
  "reviewerName": "Abbey",
  "reviewText": "I now have 4 of the 5 available colors of this shirt... ",
  "summary": "Comfy, flattering, discreet--highly recommended!",
  "unixReviewTime": 1514764800
}
```

Pipeline Architecture Diagram

18



Pipeline Architecture Diagram

19

เส้นทางการไหลของข้อมูล (Data Flow)



ระบุ Trigger สำหรับ Retraining:

"เมื่อมีข้อมูลใหม่เกิน 10,000 รีวิว → เริ่ม Retrain อัตโนมัติ"

Pipeline Architecture Plan

20

Diagram นี้ครอบคลุม

- ข้อมูลจาก Amazon Reviews → เก็บใน S3
- ประมวลผลด้วย AWS Glue
- เก็บผลลัพธ์ลง S3 (Processed)
- ฝึกโมเดล LLM บน SageMaker
- Deploy ผ่าน SageMaker Endpoint
- เรียกใช้ผ่าน API Gateway / Lambda
- แสดงผลผ่าน Dashboard (Streamlit / QuickSight)

Pipeline Architecture Plan

21

1. Data Ingestion (การเก็บรวบรวมข้อมูล)

- ข้อมูลรีวิวจาก Amazon Customer Reviews Dataset
- อัปโหลดแบบ batch เข้าสู่ Amazon S3
- จำลองการ stream ข้อมูลใหม่โดยแบ่งตามช่วงเวลา (reviewTime)
- หากต่อ ยอดได้ อาจใช้ AWS Lambda หรือ AWS Kinesis ในอนาคต

Pipeline Architecture Plan

22

2. Data Preparation & Processing (การประมวลผลและเตรียมข้อมูล)

- ใช้ AWS Glue / PySpark ทำ ETL:
 - ล้างข้อมูล (เช่น ตัดข้อความที่สั้นเกินไป / ไม่มี rating)
 - สร้าง label sentiment จากคะแนน (overall)
 - แปลงข้อมูลเป็นรูปแบบที่ใช้กับโมเดล เช่น JSONL
- จัดการข้อมูลให้พร้อมใช้งานกับ SageMaker (training-ready)

Pipeline Architecture Plan

23

3. Data Storage (การจัดเก็บข้อมูล)

- จัดเก็บข้อมูลดิบและข้อมูลหลัง ETL ไว้ที่ Amazon S3
 - Raw Zone → ข้อมูลต้นฉบับ
 - Processed Zone → ข้อมูลพร้อมเทรน
- เก็บ โมเดลที่ฝึกแล้ว (Model Artifact) บน S3 เพื่อเรียกใช้ภายหลัง

Pipeline Architecture Plan

24

4. Model Training & Deployment (การฝึกและปรับใช้โมเดล)

- ใช้ Amazon SageMaker สำหรับ:
 - Fine-tune LLM เช่น DistilBERT, T5 บนข้อมูล sentiment
 - ตั้งรอบฝึกอัตโนมัติ (retrain) เมื่อมีข้อมูลใหม่
- Deploy ผ่าน SageMaker Endpoint หรือ AWS Lambda + API Gateway

5. Visualization & Monitoring (การแสดงผลและติดตามระบบ)

- ใช้ Streamlit หรือ Amazon QuickSight สร้าง Dashboard แสดง:
 - Accuracy / F1-score ของโมเดลแต่ละรุ่น
 - แนวโน้ม Sentiment ของลูกค้า (รายวัน/รายเดือน)
 - ความเร็วในการตอบกลับ (Latency)
 - เวอร์ชันของโมเดลที่ถูกใช้งานล่าสุด

โมเดล LLM ที่คาดว่าจะได้หลังจากจัดทำ คือ “ระบบอัจฉริยะ” ที่สามารถ

- วิเคราะห์ความรู้สึก
- สรุปข้อความ
- ตอบคำถามจากข้อความ
- อัปเดตตัวเองต่อเนื่องจากข้อมูลใหม่

ผลลัพธ์ทั้งหมดนี้ → สามารถนำไปใช้ตัดสินใจทางธุรกิจ / พัฒนาองค์กร / ประหยัดแรง

อติปรายในสิ่งที่ได้เรียนรู้และแนวทางในการพัฒนาต่อยอด



27







- บทเรียนสำคัญ:
 - ความท้าทายในการจัดการข้อมูล Streaming
 - การปรับ Hyperparameters ของ SageMaker
 - การศึกษาเพิ่มเติม ใน Model ที่ใช้งานและการใช้งาน AWS
- แนวทางพัฒนาต่อ:
 - เพิ่ม Feature เช่น Summarization (ด้วย T5)
 - เพิ่ม Model ที่ใช้สามารถใช้ในการเปรียบเทียบให้ครอบคลุม

GitHub Repo


28

<https://github.com/taressup/bigdata>

 taressup / bigdata




[Code](#) [Issues](#) [Pull requests](#) [Actions](#) [Projects](#) [Wiki](#) [Security](#) [Insights](#) [Settings](#)

 **bigdata** Public





[Pin](#) [Watch](#) 0 [Fork](#) 0 [Star](#) 0

[main](#) [1 Branch](#) [0 Tags](#)


[Add file](#) [Code](#)

 **taress** Code Project

6f246a2 · 8 hours ago 1 Commit

 Cluster.py	Code Project	8 hours ago
 apigeteway.py	Code Project	8 hours ago
 etljsontocsv.py	Code Project	8 hours ago
 index.html	Code Project	8 hours ago

[README](#)



About

No description, website, or topics provided.

[Activity](#)

0 stars

0 watching

0 forks

Releases

No releases published

[Create a new release](#)

Packages

No packages published

[Publish your first package](#)

ทดลองใน **AWS**

AWS Project

30

อันนี้ก็จะเป็นการสร้าง S3 bucket ขึ้นมา จากนั้นภายในจะมีอยู่ 3 โฟลเดอร์นะครั้นมีไฟล์ ที่ชื่อว่า Raw แล้วก็มี process แล้วก็มี ตัว Cluster

The screenshot shows the AWS S3 console's 'Upload' page. The breadcrumb navigation at the top indicates the path: Amazon S3 > Buckets > s3-0120 > raw/ > Upload. The main heading is 'Upload' with an 'Info' link. Below this, a message states: 'Add the files and folders you want to upload to S3. To upload a file larger than 160GB, use the AWS CLI, AWS SDKs or Amazon S3 REST API. [Learn more](#)'. A dashed box contains the instruction: 'Drag and drop files and folders you want to upload here, or choose **Add files** or **Add folder**.' Below this, a section titled 'Files and folders (1 total, 6.1 MB)' shows a table of items to be uploaded. The table has columns for Name, Folder, Type, and Size. One item, 'small.json', is listed with a folder of '-' and a type of 'application/json' and size of '6.1 MB'. To the right of the table are buttons for 'Remove', 'Add files', and 'Add folder'. Below the table, the 'Destination' section is expanded, showing the path 's3://s3-0120/raw/'. Under 'Destination details', it says 'Bucket settings that impact new objects stored in the specified destination.' Further down, 'Permissions' and 'Properties' sections are partially visible. At the bottom right, there are 'Cancel' and 'Upload' buttons. The footer of the console shows 'CloudShell', 'Feedback', and copyright information for Amazon Web Services, Inc.

aws Search [Option+S]

United States (N. Virginia) voclabs/user3787105=tares.supa@dome.tu.ac.th @ 1544-9559-5455

Amazon S3 > Buckets > s3-0120 > raw/ > Upload

Upload [Info](#)

Add the files and folders you want to upload to S3. To upload a file larger than 160GB, use the AWS CLI, AWS SDKs or Amazon S3 REST API. [Learn more](#)

Drag and drop files and folders you want to upload here, or choose **Add files** or **Add folder**.

Files and folders (1 total, 6.1 MB) [Remove](#) [Add files](#) [Add folder](#)

All files and folders in this table will be uploaded.

<input type="checkbox"/>	Name	Folder	Type	Size
<input type="checkbox"/>	small.json	-	application/json	6.1 MB

Destination [Info](#)

Destination
[s3://s3-0120/raw/](#)

► Destination details
Bucket settings that impact new objects stored in the specified destination.

► Permissions
Grant public access and access to other AWS accounts.

► Properties
Specify storage class, encryption settings, tags, and more.

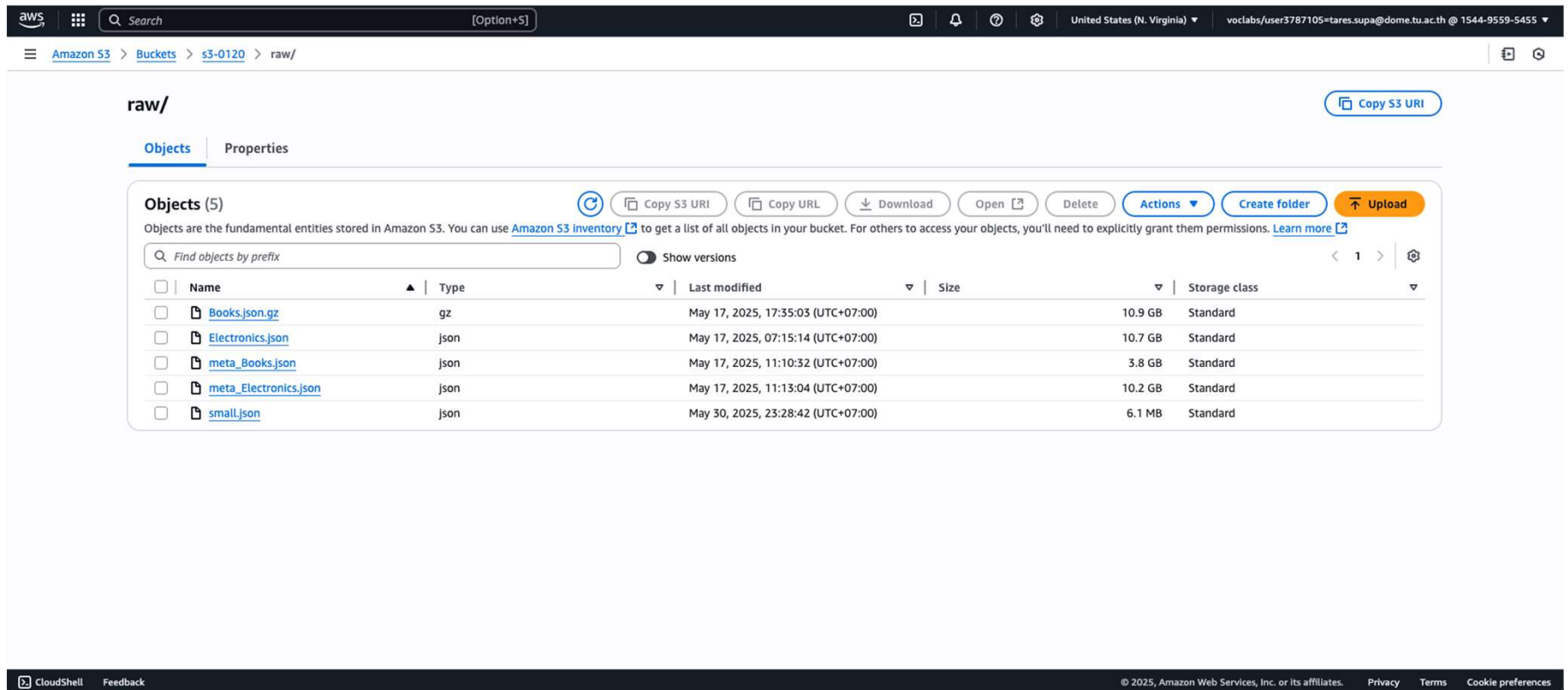
[Cancel](#) [Upload](#)

CloudShell Feedback © 2025, Amazon Web Services, Inc. or its affiliates. [Privacy](#) [Terms](#) [Cookie preferences](#)

AWS Project

31

อย่างแรกก็เนื่องจากเป็นข้อมูลดิบหลักที่ใช้ในการที่จะนำเข้าไป process ก็จะทำให้ Json ที่เป็นข้อมูลมาใส่ในไฟล์ Raw



The screenshot displays the AWS S3 console interface for a bucket named 'raw/'. The breadcrumb navigation shows 'Amazon S3 > Buckets > s3-0120 > raw/'. The 'raw/' bucket is selected, and the 'Objects' tab is active. A 'Copy S3 URI' button is visible in the top right. Below the bucket name, there are tabs for 'Objects' and 'Properties'. The 'Objects' tab shows a list of 5 objects. Above the list, there are action buttons: 'Copy S3 URI', 'Copy URL', 'Download', 'Open', 'Delete', 'Actions', 'Create folder', and 'Upload'. A search bar with the placeholder 'Find objects by prefix' and a 'Show versions' toggle are also present. The object list has columns for Name, Type, Last modified, Size, and Storage class.

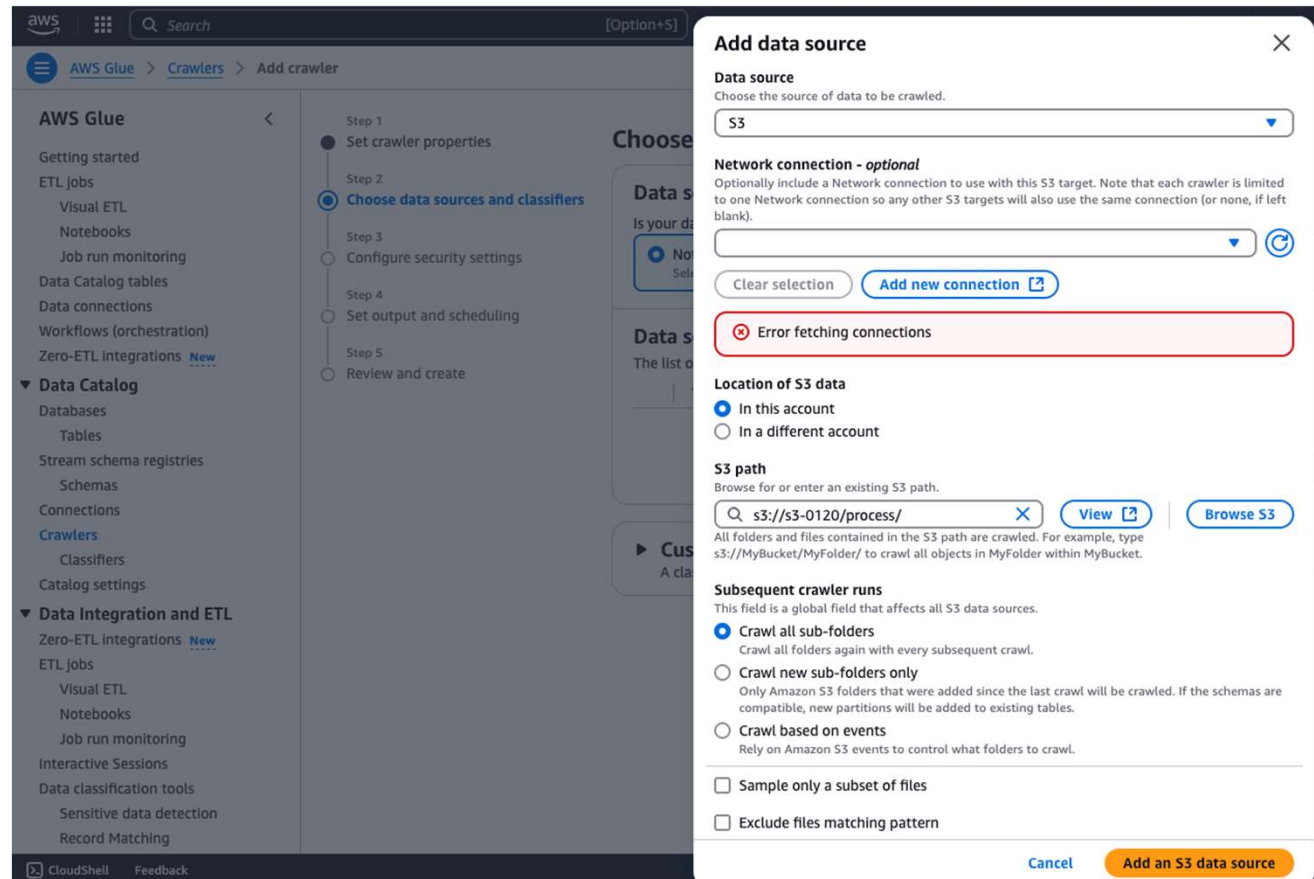
Name	Type	Last modified	Size	Storage class
Books.json.gz	gz	May 17, 2025, 17:35:03 (UTC+07:00)	10.9 GB	Standard
Electronics.json	json	May 17, 2025, 07:15:14 (UTC+07:00)	10.7 GB	Standard
meta_Books.json	json	May 17, 2025, 11:10:32 (UTC+07:00)	3.8 GB	Standard
meta_Electronics.json	json	May 17, 2025, 11:13:04 (UTC+07:00)	10.2 GB	Standard
small.json	json	May 30, 2025, 23:28:42 (UTC+07:00)	6.1 MB	Standard

© 2025, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences

AWS Project

32

ต่อมาจะเป็นขั้นตอนการสร้างตัว
aws glue ซึ่งจะต้องตั้งชื่อตัว
Crawler โดยการสร้าง Crawler
ขึ้นมา 1 ตัวจากนั้นก็ทำการ Add
Data Source เพื่อเอาข้อมูลจากใน
S3 bucket ที่ชื่อในโฟลเดอร์ที่ชื่อว่า
etl



AWS Project

33

จากนั้นก็เลือก I am role เพื่อใช้สิทธิ์ตามที่ learner Lab มีให้

The screenshot shows the AWS Glue console interface during the 'Add crawler' process. The left sidebar contains the navigation menu with categories like 'AWS Glue', 'Data Catalog', and 'Data Integration and ETL'. The main content area is titled 'Configure security settings' and shows a progress bar with five steps: 'Set crawler properties', 'Choose data sources and classifiers', 'Configure security settings' (current step), 'Set output and scheduling', and 'Review and create'. The 'IAM role' section shows an 'Existing IAM role' dropdown set to 'voclabs', with buttons for 'Create new IAM role' and 'Update chosen IAM role'. Below this is the 'Lake Formation configuration - optional' section with a checkbox for 'Use Lake Formation credentials for crawling S3 data source'. At the bottom is the 'Security configuration - optional' section with a checkbox for 'Enable at-rest encryption with a security configuration'. Navigation buttons 'Cancel', 'Previous', and 'Next' are at the bottom right.

aws [Search] [Option+S] United States (N. Virginia) voclabs/user3787105=tares.supa@dome.tu.ac.th @ 1544-9559-5455

AWS Glue > Crawlers > Add crawler

AWS Glue

- Getting started
- ETL Jobs
 - Visual ETL
 - Notebooks
 - Job run monitoring
- Data Catalog tables
- Data connections
- Workflows (orchestration)
- Zero-ETL Integrations **New**
- ▼ **Data Catalog**
 - Databases
 - Tables
 - Stream schema registries
 - Schemas
 - Connections
 - Crawlers
 - Classifiers
 - Catalog settings
- ▼ **Data Integration and ETL**
 - Zero-ETL Integrations **New**
 - ETL Jobs
 - Visual ETL
 - Notebooks
 - Job run monitoring
 - Interactive Sessions
 - Data classification tools
 - Sensitive data detection
 - Record Matching

Step 1: Set crawler properties
Step 2: Choose data sources and classifiers
Step 3: **Configure security settings**
Step 4: Set output and scheduling
Step 5: Review and create

Configure security settings

IAM role [Info](#)

Existing IAM role

voclabs [View](#)

[Create new IAM role](#) [Update chosen IAM role](#)

Only IAM roles created by the AWS Glue console and have the prefix "AWSGlueServiceRole-" can be updated.

Lake Formation configuration - optional

Allow the crawler to use Lake Formation credentials for crawling the data source. [Learn more.](#)

☐ Use Lake Formation credentials for crawling S3 data source

Checking this box will allow the crawler to use Lake Formation credentials for crawling the data source. If the data source is registered in another account, you must provide the registered account ID. Otherwise, the crawler will crawl only those data sources associated to the account. Only applicable to S3, Glue Catalog, Iceberg, and Hudi data sources.

► **Security configuration - optional**

Enable at-rest encryption with a security configuration.

[Cancel](#) [Previous](#) [Next](#)

CloudShell Feedback © 2025, Amazon Web Services, Inc. or its affiliates. [Privacy](#) [Terms](#) [Cookie preferences](#)

AWS Project

34

จากนั้นก็จะต้องตั้งค่า database ซึ่งตอนนี้ยังไม่มีนะ ก็เลยต้องไปสร้าง database ใน AWS glue ในตัว AWS glue จะมี Data catalog ขึ้นมา ก็ตั้งชื่อไฟล์นามชื่อ electronic_db

The screenshot shows the AWS Glue console interface. The top navigation bar includes the AWS logo, a search bar, and the user's profile. The left sidebar shows the navigation menu with 'AWS Glue' selected, and 'Data Catalog' expanded to show 'Databases'. The main content area is titled 'Databases (3)' and includes a description: 'A database is a set of associated table definitions, organized into a logical group.' Below this is a search bar labeled 'Filter databases'. A table lists the databases with columns for Name, Description, Location URI, and Created on (UTC). The table contains three entries: 'default', 'electronic_db', and 'products_0120_xvwu'. The 'electronic_db' entry is highlighted.

<input type="checkbox"/>	Name	Description	Location URI	Created on (UTC)
<input type="checkbox"/>	default	-	-	May 24, 2025 at 02:33:22
<input type="checkbox"/>	electronic_db	-	-	May 30, 2025 at 16:37:08
<input type="checkbox"/>	products_0120_xvwu	-	-	May 24, 2025 at 02:21:46

AWS Project

35

จากนั้นให้ทำการสร้าง API Gateway ตั้งชื่อ API ขึ้นมา แล้วก็ทำการสร้าง route เพื่อให้มันรู้ที่อยู่ของการดึงข้อมูลหรือว่าผ่านทางออกไปใช้งาน

The screenshot shows the AWS Management Console interface for editing an API Gateway route. The top navigation bar includes the AWS logo, a search bar, and user information. The left sidebar shows the 'API Gateway' section with options like 'APIs', 'Custom domain names', 'Domain name access associations', and 'VPC links'. The main content area is titled 'Edit route' and features a green notification banner at the top stating 'Successfully updated 1 routes.' Below this, the 'Route and method' section is visible, showing the 'Route name eg. /pets' and a dropdown menu for the method set to 'GET'. The path '/get-data' is entered in the adjacent text field. At the bottom right of the form, there are 'Cancel' and 'Save' buttons.

AWS Project

36

ต่อไปเป็นตัวอย่าง integration เอาไว้สำหรับดึงข้อมูล จากตัวโดยใช้ต้องไป Trigger กับตัว Lambda เพื่อเอาข้อมูลในนั้นนำมาใช้งาน

The screenshot shows the AWS Management Console interface for editing an API Gateway integration. The breadcrumb navigation at the top indicates the path: **API Gateway** > **APIs** > **sentiment (g8e8fe7sel)** > **Integrations**.

API Gateway

- APIs
- Custom domain names
- Domain name access associations
- VPC links

API: sentiment...(g8e8fe7sel)

Develop

- Routes
- Authorization
- Integrations**
- CORS
- Reimport
- Export

Deploy

Successfully updated 1 routes.

Notifications 0 0 2 0 0

Edit integration

Integration target

Integration type
Lambda function
Send the request to a Lambda function.

Integration details

Integration target
Choose the Lambda function that API Gateway invokes when the route receives a request.

AWS Region
us-east-1

Lambda function
arn:aws:lambda:us-east-1:673184257118:function:getData

Advanced settings

จากนั้นก็ให้ **Query** ข้อมูลตัวนี้ก็จะเป็นตัวที่เอาไว้สำหรับการดึงข้อมูล ไม่ตั้งค่ามันก็จะติด **permission denied** ก็คือไม่สามารถนำข้อมูลนี้ไปใช้งานได้ ก็จะติด **Core** สถานะมันก็จะ เป็น **500** ก็คือขึ้นแสดง **Status Error** ขึ้นมา ก็ต้องตั้งค่าโดยการสั่งให้ปลดล็อกทั้งหมดโดยการ * แล้วก็เลือก **Method** เป็น **GET** นะครับ

API Gateway > APIs > sentiment (g8e8fe7sel) > CORS

API Gateway

- APIs
- Custom domain names
- Domain name access associations
- VPC links

API: sentiment...(g8e8fe7sel)

▼ Develop

- Routes
- Authorization
- Integrations
- CORS**
- Reimport
- Export

▼ Deploy

- Stages

▼ Monitor

Successfully updated 1 routes.

Notifications 0 0 2 0 0 0

Cross-Origin Resource Sharing

Stage: - Deploy

Configure CORS [Info](#)

Configure Clear

CORS allows resources from different domains to be loaded by browsers. If you configure CORS for an API, API Gateway ignores CORS headers returned from your backend integration. See our [CORS documentation](#) for more details.

Access-Control-Allow-Origin	Access-Control-Allow-Headers
* X	No headers are allowed
Access-Control-Allow-Methods	Access-Control-Expose-Headers
GET X	No Expose Headers are allowed
Access-Control-Max-Age	Access-Control-Allow-Credentials
0 Seconds	NO

AWS Project

38

จากนั้นก็ทำการสร้าง Stage ขึ้นมา 1 ตัว ก็อันนี้จะเป็นการจำลองโดยใช้ บริการสร้าง state ที่ชื่อว่า Dev ขึ้นมาเอาไว้สำหรับการเป็นตัวกลางในการเรียกข้อมูล

The screenshot shows the AWS API Gateway console. The breadcrumb navigation is: API Gateway > APIs > sentiment (g8e8fe7sel) > Stages. A green banner at the top indicates 'Successfully updated 1 routes.' Below this, the 'Stages' section for the 'sentiment' API is shown. It includes a 'Create' button and a search bar. Two stages are listed: '\$default' and 'dev' (selected). The 'dev' stage details are expanded, showing its name, creation date (May 31, 2025 1:23 AM), invoke URL, and deployment status (Disabled). The 'Stage variables' section is also visible at the bottom.

API Gateway > APIs > sentiment (g8e8fe7sel) > Stages

API: sentiment...(g8e8fe7sel)

Develop

- Routes
- Authorization
- Integrations
- CORS
- Reimport
- Export

Deploy

- Stages

Monitor

- Metrics
- Logging

Protect

Successfully updated 1 routes.

Notifications 0 0 2 0 0 0

Stages

Stages for sentiment [Create](#)

Search

☐ \$default

☒ dev

Stage details

[Delete](#) [Edit](#)

Details

Name	Created	Last updated
dev	May 31, 2025 1:23 AM	May 31, 2025 1:23 AM

Invoke URL

<https://g8e8fe7sel.execute-api.us-east-1.amazonaws.com/dev>

Description

None

Attached deployment

Automatic Deployment

☐ Disabled

Stage variables

Find resources

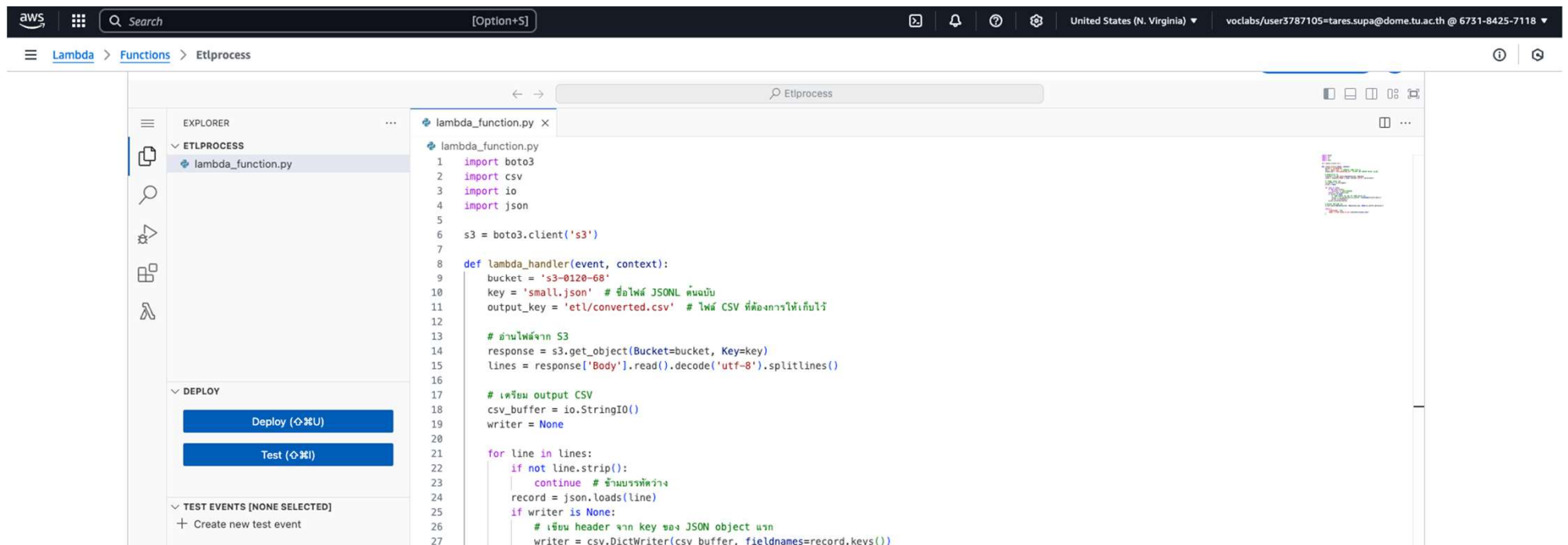
< 1 >

Key	Value
-----	-------

AWS Project Process

39

ต่อไปจะเป็นขบวนการ process ก็จะเริ่มต้นจากการสร้าง Lambda ฟังก์ชันขึ้น แล้วก็ทำการตั้งชื่อว่า etlprocess ซึ่งจะเป็นกระบวนการในการแปลงข้อมูลจาก json ไฟล์นะครับเป็น csv เนื่องจากว่า การใช้ csv มันจะทำให้มีการประมวลผลที่เร็วขึ้นจาก json ความเร็วก็จะเพิ่มขึ้นประมาณ 40 ถึง 60 เปอร์เซ็นต์



The screenshot shows the AWS Lambda console interface. The breadcrumb navigation at the top indicates the path: Lambda > Functions > Etlprocess. The left sidebar contains the 'EXPLORER' panel with a tree view showing 'ETLPROCESS' and 'lambda_function.py'. Below this is the 'DEPLOY' section with 'Deploy' and 'Test' buttons, and a 'TEST EVENTS' section with a 'Create new test event' button. The main area displays the Python code for 'lambda_function.py'. The code imports boto3, csv, io, and json, initializes an S3 client, and defines a lambda_handler function that retrieves a JSON object from S3, converts it to CSV, and writes it to S3. The code is as follows:

```
1 import boto3
2 import csv
3 import io
4 import json
5
6 s3 = boto3.client('s3')
7
8 def lambda_handler(event, context):
9     bucket = 's3-0120-68'
10    key = 'small.json' # ชื่อไฟล์ JSONL ต้นฉบับ
11    output_key = 'etl/converted.csv' # ไฟล์ CSV ที่ต้องการให้เก็บไว้
12
13    # อ่านไฟล์จาก S3
14    response = s3.get_object(Bucket=bucket, Key=key)
15    lines = response['Body'].read().decode('utf-8').splitlines()
16
17    # เตรียม output CSV
18    csv_buffer = io.StringIO()
19    writer = None
20
21    for line in lines:
22        if not line.strip():
23            continue # ข้ามบรรทัดว่าง
24        record = json.loads(line)
25        if writer is None:
26            # เขียน header จาก key ของ JSON object แรก
27            writer = csv.DictWriter(csv_buffer, fieldnames=record.keys())
```

AWS Project Process

40

คราวนี้เมื่อลองจำลองในการ Test Lambda ก็ถ้าผลมันถูกต้องก็จะขึ้น Status 200 ตามรูปนี้ ซึ่งมันจะแปลงเป็น save File เรียบร้อยแล้ว

The screenshot displays the AWS Lambda console interface. The top navigation bar shows the user is logged in as 'voclabs/user3787105=tares.supa@dome.tu.ac.th' in the 'United States (N. Virginia)' region. The main content area is divided into several sections:

- EXPLORER:** Shows the file structure with 'GETDATAETL' and 'lambda_function.py'.
- DEPLOY:** Contains buttons for 'Deploy (Ctrl+U)' and 'Test (Ctrl+I)'.
- TEST EVENTS [SELECTED: TEST-S3-READ]:** Lists test events, including 'test-s3-read'.
- ENVIRONMENT VARIABLES:** A section for managing environment variables.
- Execution Results:** The central pane shows the output of the function execution. It indicates 'Status: Succeeded' and 'Test Event Name: test-s3-read'. The response is a JSON object:

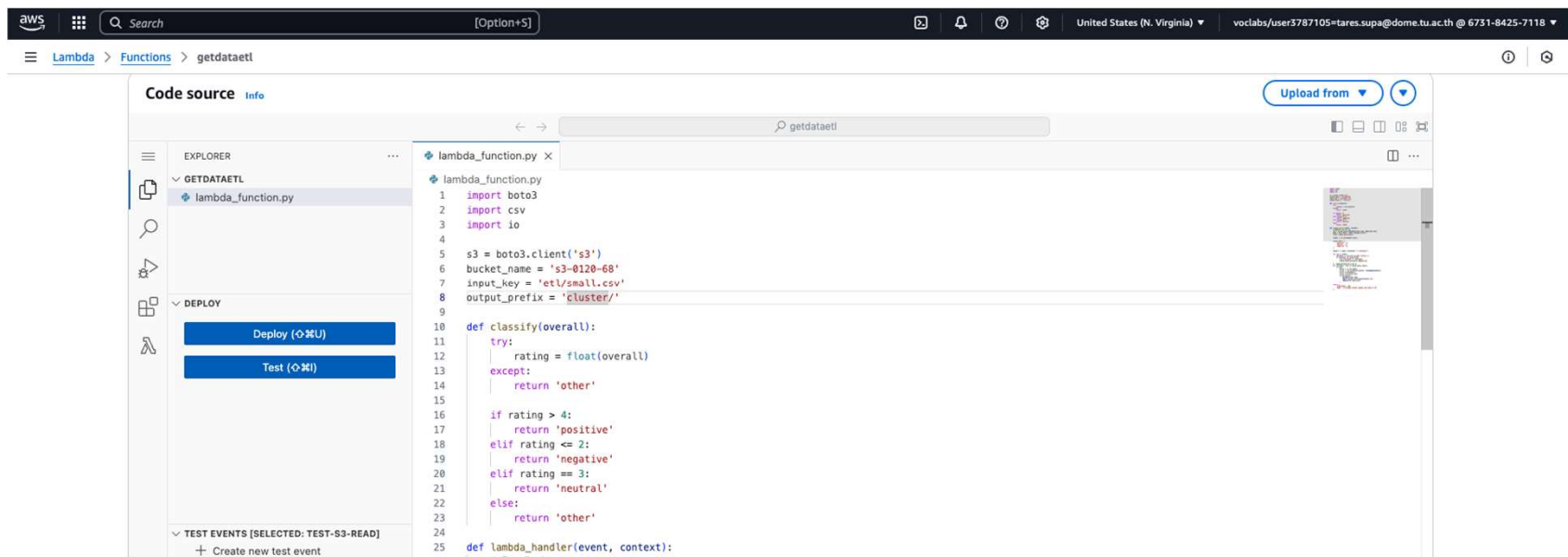
```
{  "statusCode": 200,  "body": "✅ Grouped without pandas and saved to S3"}
```

. Below this, the 'Function Logs' section provides detailed execution metrics:
 - START RequestId: 6e4fe3fe-d67a-46c3-9167-0b954c4aac72 Version: \$LATEST
 - END RequestId: 6e4fe3fe-d67a-46c3-9167-0b954c4aac72
 - REPORT RequestId: 6e4fe3fe-d67a-46c3-9167-0b954c4aac72 Duration: 3242.85 ms Billed Duration: 3243 ms Memory Size: 128 MB Max Memory Used: 101 MB Init Duration: 392.76 ms

AWS Project Process

41

ต่อมาก็จะเป็นตัวสร้าง Lambda ตัวนี้ขึ้นมาที่ชื่อว่าตัว getdataetl ก็ตัว Lambda ตัวนี้ก็จะเอาไว้สำหรับทำการ cluster คือจะเอาตัว etl ที่ผ่านการประมวลผลแล้วเป็นไฟล์ csv เอามาทำ cluster จำแนกเป็นกลุ่ม แล้วก็ โดยใช้ Lambda แบ่งเป็น 3 กลุ่ม ถ้ามีกลุ่มที่เป็น Positive Natural



```
1 import boto3
2 import csv
3 import io
4
5 s3 = boto3.client('s3')
6 bucket_name = 's3-0120-68'
7 input_key = 'etl/small.csv'
8 output_prefix = 'cluster/'
9
10 def classify(overall):
11     try:
12         rating = float(overall)
13     except:
14         return 'other'
15
16     if rating > 4:
17         return 'positive'
18     elif rating <= 2:
19         return 'negative'
20     elif rating == 3:
21         return 'neutral'
22     else:
23         return 'other'
24
25 def lambda_handler(event, context):
```

AWS Project Process

42

ดังนั้นก็จะทำการตั้งค่า ตัว trigger ตัว trigger นี้ก็เอาไว้สำหรับการ ให้ trigger ข้อมูลแบบ Auto เมื่อมีข้อมูลอยู่ในตัว S3 bucket เรียบร้อยแล้ว

aws

Search

[Option+S]

United States (N. Virginia)

voclabs/user3787105=tares.supa@dome.tu.ac.th @ 6731-8425-7118

Lambda

Functions


Etlprocess

Edit trigger

Info


Edit trigger

Trigger configuration

 S3
aws asynchronous storage

Bucket

Choose or enter the ARN of an S3 bucket that serves as the event source. The bucket must be in the same region as the function.



Bucket must be in region us-east-1

Event types

Select the events that you want to have trigger the Lambda function. You can optionally set up a prefix or suffix for an event. However, for each bucket, individual events cannot have multiple configurations with overlapping prefixes or suffixes that could match the same object key.

PUT

X

Prefix - optional

Enter a single optional prefix to limit the notifications to objects with keys that start with matching characters. Any [special characters](#) must be URL encoded.

Suffix - optional

Enter a single optional suffix to limit the notifications to objects with keys that end with matching characters. Any [special characters](#) must be URL encoded.

Lambda will add the necessary permissions for AWS S3 to invoke your Lambda function from this trigger. [Learn more](#) about the Lambda permissions model.

Cancel

Save

AWS Project Process

43

ดึงข้อมูลมาใช้งานโดยใน AWS Athena

The screenshot displays the AWS Athena Query Editor interface. The top navigation bar includes the AWS logo, a search bar, and user information. The main interface is divided into several sections:

- Left Sidebar:** Contains the 'Data' section with dropdowns for 'Data source' (AwsDataCatalog), 'Catalog' (None), and 'Database' (electronic_db). Below this is the 'Tables and views' section with a 'Create' button and a search bar for filtering tables and views. A list of tables is shown, including 'etl'.
- Query Editor:** The central area where the SQL query is written. The query is: `SELECT * FROM "electronic_db"."etl";`. Below the query editor are buttons for 'Run again', 'Explain', 'Cancel', 'Clear', and 'Create'.
- Query Results:** The bottom section shows the results of the query. It indicates the query is 'Completed' with a green checkmark. Performance metrics are displayed: 'Time in queue: 107 ms', 'Run time: 941 ms', and 'Data scanned: 2.86 MB'. There are buttons for 'Copy' and 'Download results CSV'. Below the metrics is a table of results.

Query Results Table:

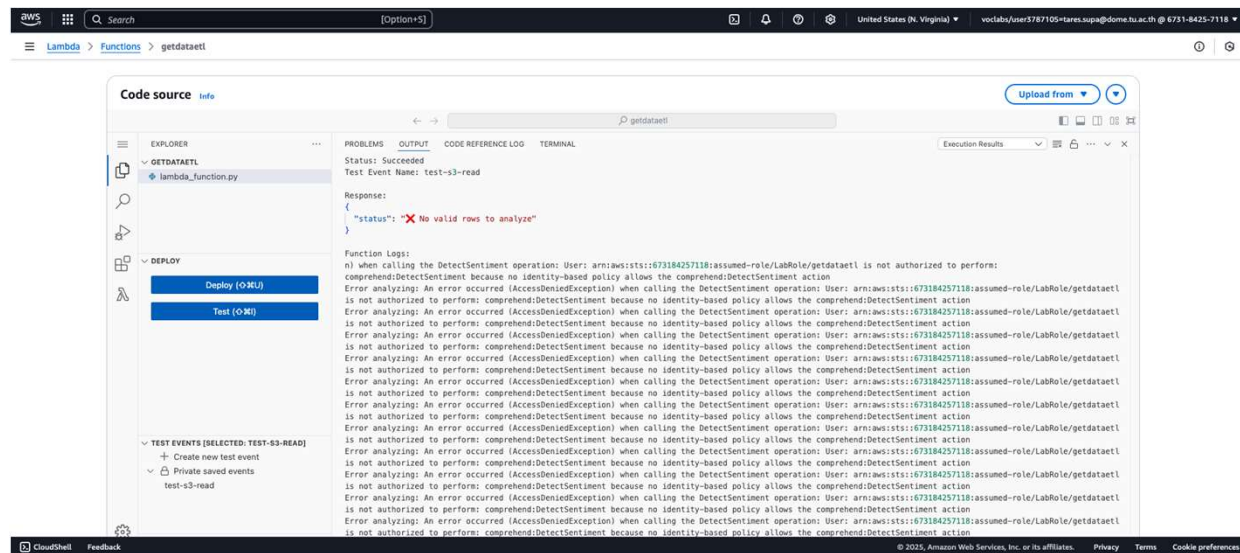
#	overall	verified	reviewtime
1	5.0	true	"07 17
2			1026864000
3	5.0	false	"07 6
4	5.0	false	"06 28

The footer of the interface includes links for 'CloudShell', 'Feedback', and copyright information for Amazon Web Services, Inc. or its affiliates, along with links for 'Privacy', 'Terms', and 'Cookie preferences'.

AWS Project Error

44

เนื่องจากว่าไฟล์ข้อมูลมีปริมาณมหาศาลเนื่องจากว่าในตัว S3 bucket มีไฟล์ขนาดไฟล์ที่อยู่ในเรท 10-20 GB ทำให้ประมวลผลข้อมูลเนี่ยมันช้า แล้วผลจากการใช้ตัว learner Lab มันต้องมีข้อจำกัดเรื่องเวลา ทำให้มันรันได้ประมาณ 10 นาทีหรือว่า 10 เต็มที่ 35 นาที ทำให้มัน Time Out ก่อนกำหนดขั้นตอนวิธีการแก้ปัญหาอย่างแรกเลยคือการลดขนาดไฟล์ลงนอกจากสิทธิของก็ตัดให้เหลือแค่ 10 MB แต่ก็ยังใช้เวลาเยอะอยู่ดีนะมากกว่า 10 นาทีอยู่ดีดังนั้นเราจะยังไม่ใช้ในขั้นตอนนี้



AWS Project Error

45

ทำมาก็จะใช้อีกรูปแบบหนึ่งก็คือจะใช้งานผ่านตัว AWS comprehend ซึ่งเป็นตัว aws ที่เอาไว้สำหรับประมวลผลข้อมูลขนาดใหญ่จาก aws แต่ว่าอันนี้ก็จะมีความผิดพลาดเรื่องติด permission IAM Role ทำให้ไม่สามารถใช้งานได้ ก็เลยต้องเปลี่ยนอีกตัวหนึ่งขึ้นมา

The screenshot shows the AWS console interface. At the top, there's a navigation bar with the AWS logo, a search bar, and various icons. Below the navigation bar, there's a red error banner that reads: "AccessDeniedException: User: arn:aws:sts::673184257118:assumed-role/voclabs/user3787105=tares.supa@dome.tu.ac.th is not authorized to perform: sso:ListInstances on resource: arn:aws:sso::instance/* because no identity-based policy allows the sso:ListInstances action". To the right of the error message is a button labeled "Diagnose with Amazon Q". Below the error banner, the main heading is "Create an Amazon SageMaker Unified Studio domain". Underneath this heading is a sub-heading "Choose how you want to set up your domain." followed by two options: "Quick setup (recommended for exploration)" and "Manual setup (recommended for organizations)". The "Quick setup" option is selected and highlighted with a blue border. It lists several capabilities: "Data analytics, machine learning, SQL, and generative AI", "Data and AI governance", "Generative AI app development using Amazon Bedrock serverless models", "Amazon Q - Free tier", and "Authentication via IAM or IAM Identity Center". The "Manual setup" option is also listed with its own set of capabilities. At the bottom of the page, there's a blue banner that says "No models accessible" and provides a link to "Grant model access".

AccessDeniedException: User: arn:aws:sts::673184257118:assumed-role/voclabs/user3787105=tares.supa@dome.tu.ac.th is not authorized to perform: sso:ListInstances on resource: arn:aws:sso::instance/* because no identity-based policy allows the sso:ListInstances action

Diagnose with Amazon Q

Create an Amazon SageMaker Unified Studio domain

Choose how you want to set up your domain.

How do you want to set up your domain? [Info](#)

- ☒ **Quick setup (recommended for exploration)**
Let Amazon SageMaker Unified Studio configure your domain with the following default capabilities that you customize later:
 - Data analytics, machine learning, SQL, and generative AI
 - Data and AI governance
 - Generative AI app development using Amazon Bedrock serverless models
 - Amazon Q - Free tier
 - Authentication via IAM or IAM Identity Center
- ☐ **Manual setup (recommended for organizations)**
Full control over customizing your Amazon SageMaker Unified Studio domain from scratch.
 - Customize data analytics, machine learning, SQL, generative AI and more
 - Data and AI governance
 - Configure Amazon Bedrock generative AI playgrounds and app development
 - Amazon Q - Free tier
 - Authentication via IAM, IAM Identity Center, or SAML

Existing Amazon DataZone user?
Continue using Amazon DataZone without any additional Amazon SageMaker capabilities.
[Create an Amazon DataZone domain](#)

No models accessible
Grant access to Amazon Bedrock serverless models for use in Amazon SageMaker. Explore, experiment, and collaborate using accessible models in the Generative AI playgrounds and app development tools.

[Grant model access](#)

AWS Project Error

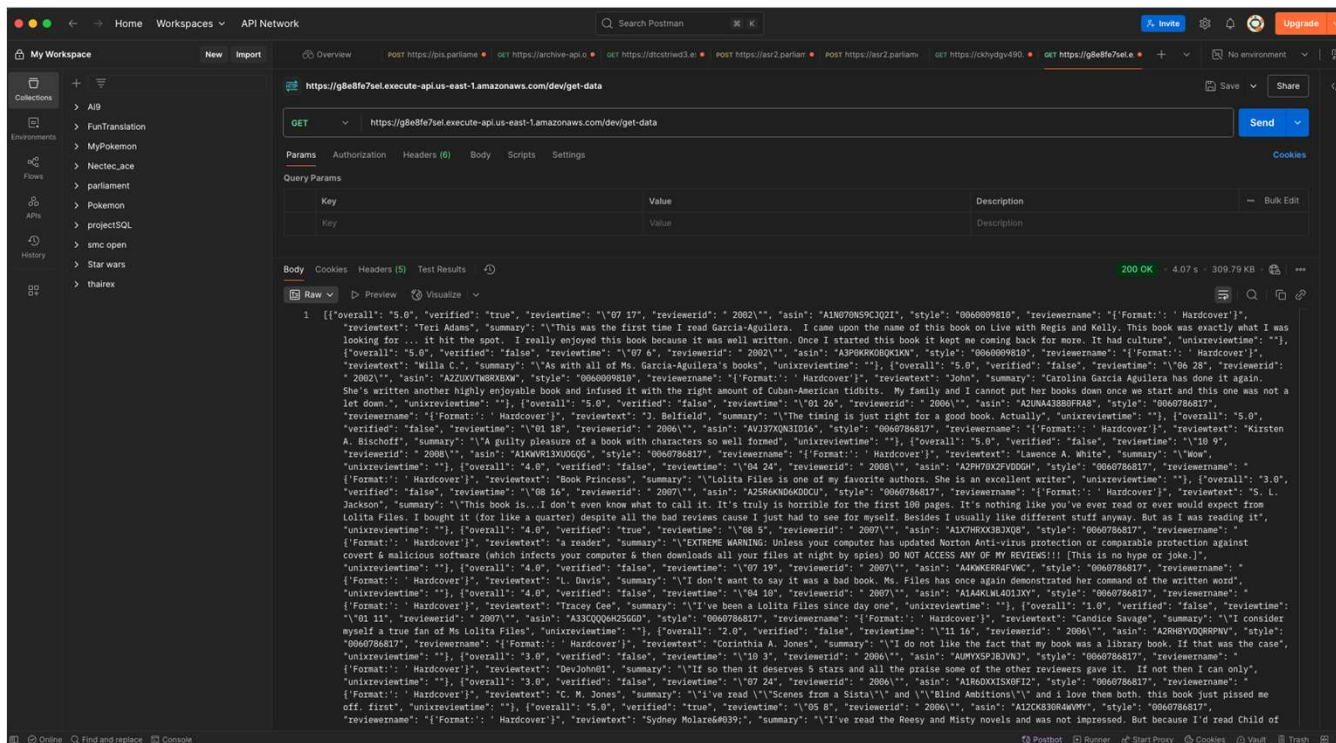
46

ก็จะใช้อีกแบบนึงนะครับก็จะเป็นตัว Quick sigh ก็เอาไว้แสดง dashboard แต่ว่าก็ติดปัญหาเรื่อง role เหมือนกัน ก็คือติด permission เรื่องปิดการเข้าถึงนั้นเนื่องจากเป็นสิทธิ์นักเรียนมันจึงไม่สามารถใช้งานได้ จึงใช้งานผ่านตัว API Gateway แทนเนื่องจากว่าเพราะว่ามันไม่ได้ก็เลยต้องใช้ API Gateway แทน เพื่อดึงข้อมูลให้ออกมาโชว์ในรูปแบบเป็นกราฟขึ้นมา

The screenshot shows the AWS console interface. At the top, there's a navigation bar with the AWS logo, a search bar, and a menu icon. Below the navigation bar, a red error banner displays the message: "AccessDeniedException: User: arn:aws:sts::673184257118:assumed-role/voclabs/user3787105=tares.supa@dome.tu.ac.th is not authorized to perform: datazone:ListDomains on resource: arn:aws:datazone:us-east-1:673184257118:domain/*". To the right of the error message is a button labeled "Diagnose with Amazon Q". Below the error banner, the main content area shows the "Amazon SageMaker" dashboard. The dashboard has a dark blue background with the text "Analytics" and "Amazon SageMaker" in white. Below this, it says "The center for data, analytics, and AI". There's a "NextGen" badge next to "Amazon SageMaker". The dashboard also includes a section titled "Get started with Amazon SageMaker Unified Studio" with a button "Create a Unified Studio domain". Another section titled "Continue with Amazon DataZone" has a button "Create an Amazon DataZone domain".

AWS Project Result

47



ทดลองการใช้ API gateway ในการดึงข้อมูลโดยใช้โปรแกรม Postman ยิง API เพื่อให้นำข้อมูลออกมาแสดงผ่านทางหน้าเว็บ

AWS Project Result

48

SET Total Value (from Athena)





THANK YOU