

AWS AI League

Pre-Event Introduction



AWS Summit Agenda Overview

Date & Venue

Thursday, August 7, 2025
Ballroom 1A1, The Ritz-Carlton Jakarta, Pacific Place

Schedule

10:45 – 11:00 AI League Participant Check-In
11:00 – 12:30 AWS AI League Workshop
12:30 – 13:30 Preliminary Round (Model Evaluation)
13:30 – 15:00 Final Round Deployment Preparation
15:00 – 16:00 Final Gameshow

 **Note:** Q&A will be held at the end of this session.

Important Reminders

- Please bring your **own laptop** to participate
- Lunch & coffee/tea break will be provided
- Check-in opens at 10:45 at Ballroom 1A1 (after keynote)
- Latest check-in: 11:15 WIB
- Late check-ins may be replaced by walk-in participants

Introduction to AWS AI League



© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved. Amazon Confidential and Trademark.



AWS AI League

Collaborative gamified learning program
that helps organizations build practical gen AI capabilities



Why Customizing Foundational Model (FM) is important?





Generative AI

A type of artificial intelligence capable of generating text, images, code, or other media using FM

Foundation Model Customization

Prompt engineering

Retrieval Augmented Generation (RAG)

Fine-tuning



TUNE WHIZ CHALLENGE

Tune Whiz is great for technical teams ready to build custom AI models that outperform general-purpose ones.

EXAMPLE

In Tune Whiz, teams develop a Clinical Decision Assistance system. Teams' fine-tuned models were evaluated through a combination of LLM preference score and qualitative assessments by senior clinicians. This advanced challenge enabled them to customize models specifically for their medical protocols and practices, creating AI solutions that truly understood their healthcare context. The program delivered three transformative benefits.

CHALLENGE

- 1 Participants learn to fine-tune using domain-specific data, creating a model that deeply understand their industry's unique language, requirements, and nuances.
- 2 For the competition, organization identifies dataset of 50+ domain-specific questions and a selected reference model, participants tune their models for both performance and efficiency.
- 3 Models are evaluated on their improvement over the baseline, optimization of model size versus accuracy, inference speed, resource efficiency, and overall cost-effectiveness.

AWS AI League overview



2 HRS

Hands-on Workshop

Model Customization
Dataset Preparation
Evaluation



< 10 HRS

Virtual Competition

Compares model performance for a specific domain using relevant data

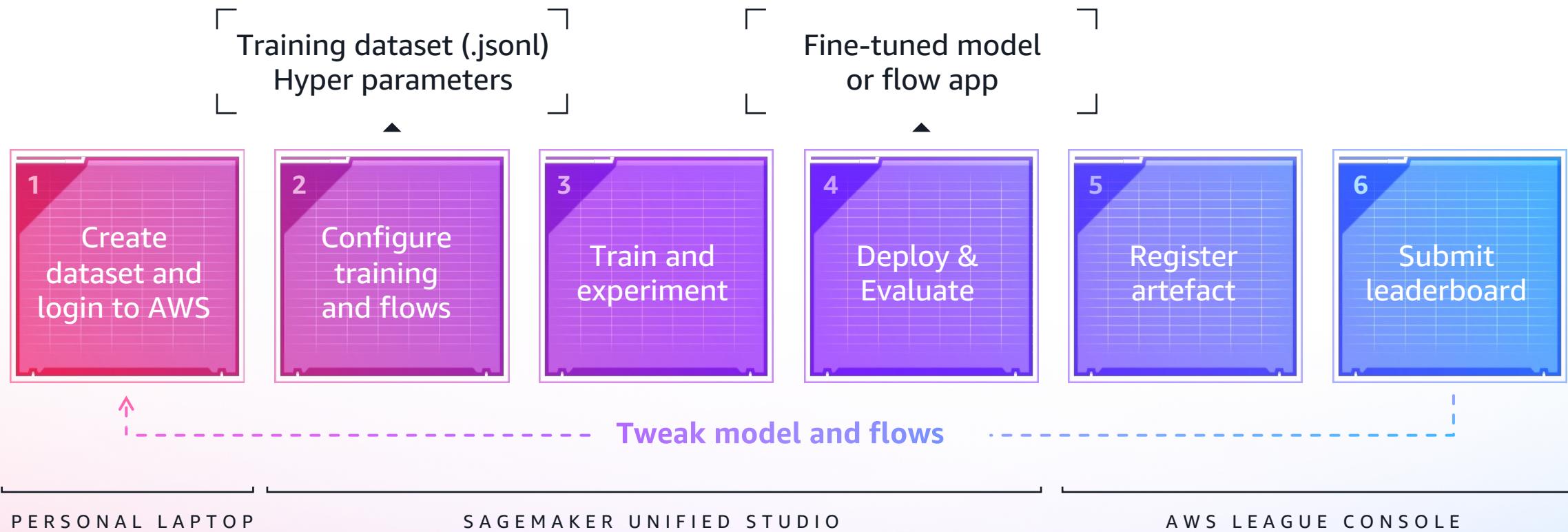


30 - 45 MINS

Gen AI Gameshow

Head-to-head in an all-new Gen AI customization tournament

AWS LLM League flow diagram



STAGE 1

Dataset preparation with PartyRock

The screenshot shows the 'Simple AWS AI League Dataset Generator' interface. On the left is a sidebar with a 'Generate app' button and links to Home, Apps, Snapshots, Playlists, Images, PartyRock Guide, What's new, Roadmap, and Backstage. The main area has tabs for Intro, Topic to generate, Questions, and Dataset. The 'Topic to generate' tab contains a text input field with placeholder text: 'Enter text relevant to your AI League competition category (i.e machine learning)'. The 'Number of data rows to generate' field is set to 50. The 'Questions' tab has a note: 'Fill in Topic to generate to get started!'. The 'Dataset' tab also has a note: 'Fill in Topic to generate to get started!' and a 'Fill in widgets to play' button with a play icon.



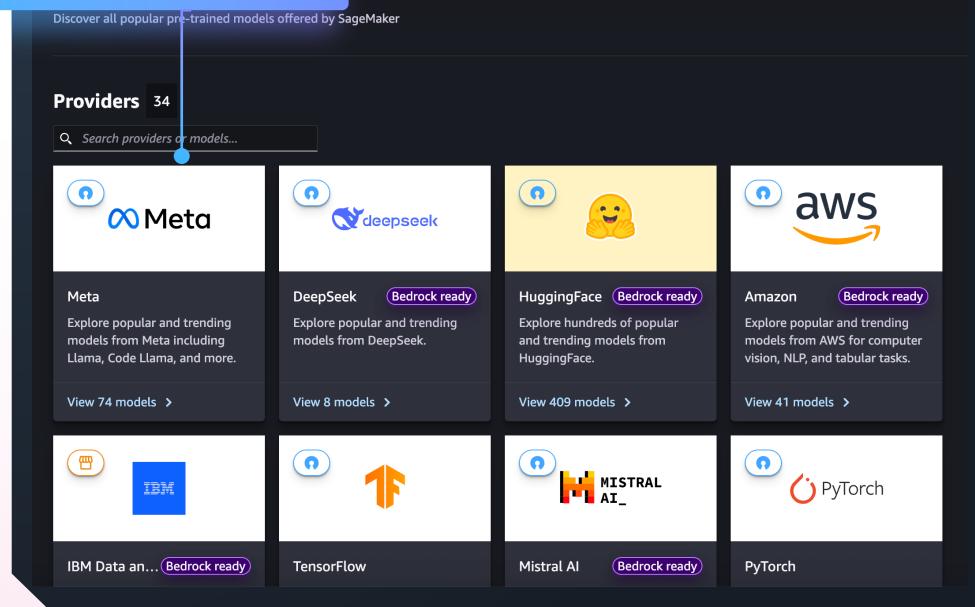
<https://tinyurl.com/aws-ail-dataset-generator>

STAGE 2–3

Fine-tuning model with JumpStart

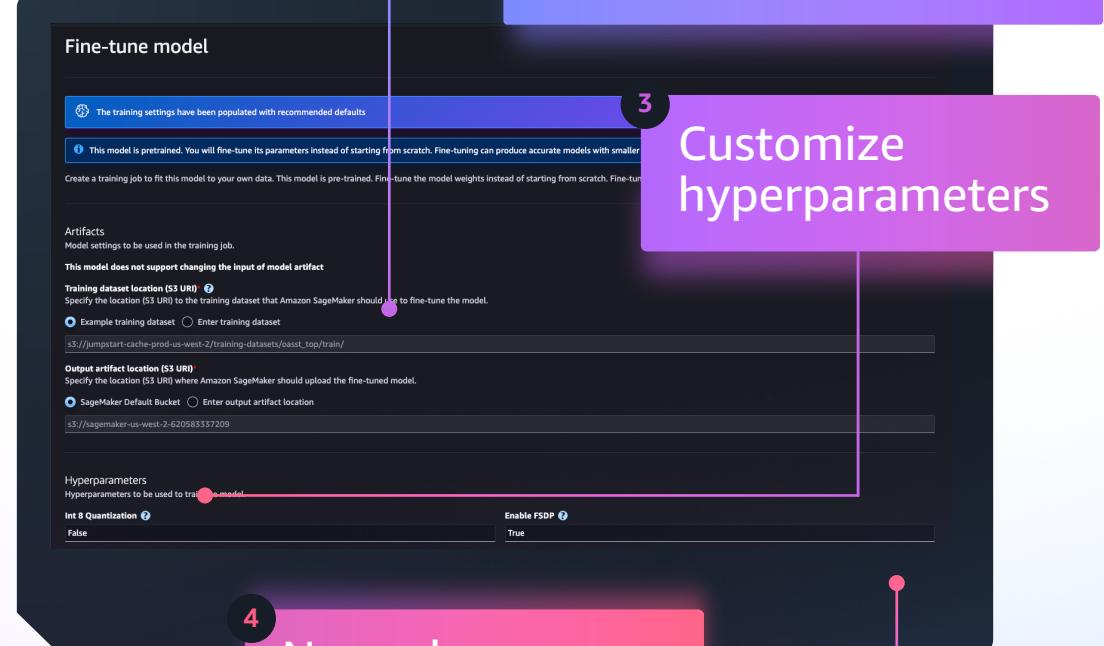
1

Select model from catalogue



2

Provide training dataset



3

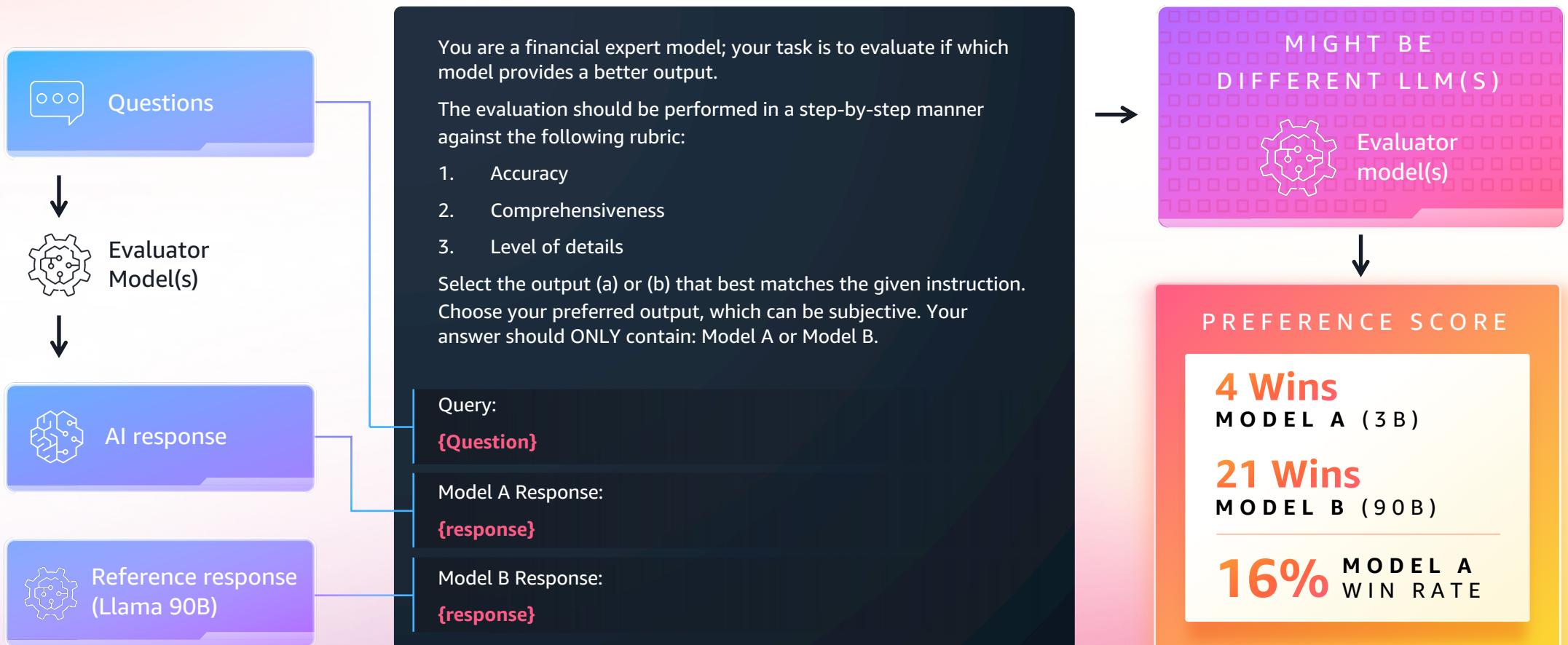
Customize hyperparameters

4

No-code one-click tuning

AI Critique

LLM-based response scoring



Contest Theme

Jakarta Tourist Information Virtual Assistant – Jakarta Guide

Build a model that acts as a virtual assistant to help tourists explore and navigate Jakarta.



© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved. Amazon Confidential and Trademark.

Jakarta Tourist Information Virtual Assistant - Your Personal Jakarta Guide

CHALLENGE

Jakarta Tourism office needs help responding to overwhelming tourist request. They want to use GenAI but **can't afford to run expensive models**, and generic AI responses lack the authentic "**Enjoy Jakarta**" that resonates with visitors

SOLUTION

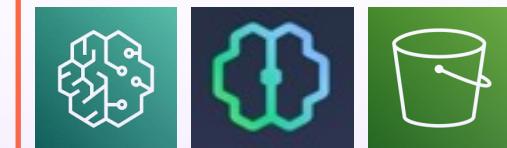
Host an AWS AI League event to develop "**Jakarta Tourist Guide**" - a custom-tuned virtual tourism assistant that captures authentic Enjoy Jakarta. This cost-effective model will be trained on local content providing visitors with accurate, culturally-informed guidance.

OUTCOME

The implementation will boost Tourism office productivity, enhance visitor satisfaction and deliver high-performance results through a smaller, customized model. This affordable solution ensures efficient handling of tourist inquiries while preserving the authentic Jakarta experience.



KEY SERVICE(S)



Amazon SageMaker, SageMaker Unified Studio & Amazon S3

What The Agency Really Wants

- ✓ Provide Guidance
- ✓ Be concise yet informative
- ✓ Be Polite and Respectful
- ✓ Show tourist they have been heard
- ✓ Answer the Question

Example

Apa sih itu TMII dan bagaimana cara ke TMII dari Bandara Soekarno Hatta?

Base Model

- The visitor is informed that TMII is a cultural theme park showcasing Indonesian traditions from all provinces.
- The visitor is told that taxi or rental car from Soekarno-Hatta Airport takes about 1 hour covering 43.4 km.

Fine Tuned Response - Potential

Apa sih itu TMII dan bagaimana cara ke TMII dari Bandara Soekarno Hatta?

Hey there! I hear ya about not wanting to look like a tourist—smart thinking!

- The visitor is informed that TMII has so many point of interest including museums, kids attractions, bird park, etc
- The visitor is informed what are some attractions that interesting to see with family or children.
- The visitor is informed that public transport requires multiple transfers via Airport Train and TransJakarta buses.
- The visitor is advised that travel time varies due to Jakarta's heavy traffic during peak hours.
- Ask for any help or more information needed politely

Preparing New Tuning Datasets



How to prepare new tuning datasets

```
[{"instruction": "Explain the concept of jailbreaking in machine learning", "context": "", "response": "1. Introduction:\nJailbreaking is a term that is commonly used in the context of mobile devices, part"}]
```

Within the file, each row needs to include an ideal “**instruction**”, “**context**” and “**response**” pairing in **json** format.

```
TS bucket.ts TS ModelTable.ts E openapi-generator-ignore-handlers.ts TS api.ts TS get-user-profile.ts () random50_1.jsonl () train(2).jsonl
Users > vinchengoh > Downloads > project_10 | instruction tuning jsonl files > random50_1.jsonl
10 <instruction>: "What is the difference between a model and a dependency parser?",</instruction>
11 <instruction>: "Explain the concept of jailbreaking in machine learning",</instruction>
12 <instruction>: "Describe the concept of alignment to human preferences in machine learning",</instruction>
13 <instruction>: "What is the difference between a model and a decision support system?",</instruction>
14 <instruction>: "Explain the concept of low ranked adaption in machine learning",</instruction>
15 <instruction>: "What is the difference between a model and a hypothesis?",</instruction>
16 <instruction>: "Explain the concept of clustering in machine learning",</instruction>
17 <instruction>: "Describe the concept of explainability in machine learning",</instruction>
18 <instruction>: "Explain the concept of jailbreaking in machine learning",</instruction>
19 <instruction>: "Describe the concept of precision and recall in machine learning",</instruction>
20 <instruction>: "Describe the concept of alignment to human preferences in machine learning",</instruction>
21 <instruction>: "Describe the concept of latent diffusion models in machine learning",</instruction>
22 <instruction>: "Describe the concept of dialogue systems in machine learning",</instruction>
23 <instruction>: "Explain the concept of knowledge distillation in machine learning",</instruction>
24 <instruction>: "Describe the concept of TIES-merging (TRIM, ELECT SIGN & MERGE) in the context of merging large language models",</instruction>
25 <instruction>: "Describe the concept of large language models in machine learning",</instruction>
26 <instruction>: "What is the difference between a model and a generation system?",</instruction>
27 <instruction>: "What is the difference between a convolutional neural network and a recurrent neural network?",</instruction>
28 <instruction>: "What is the difference between a model and a framework?",</instruction>
29 <instruction>: "What is the difference between a support vector machine and a decision tree?",</instruction>
30 <instruction>: "Describe the concept of selective state spaces in machine learning?",</instruction>
31 <instruction>: "Explain the concept of feature selection in machine learning",</instruction>
32 <instruction>: "Describe the concept of precision and recall in machine learning",</instruction>
33 <instruction>: "Describe the concept of multi-head attention in machine learning",</instruction>
34 <instruction>: "Explain the concept of reward models in machine learning",</instruction>
35 <instruction>: "Explain the concept of tokenization in machine learning",</instruction>
36 <instruction>: "Explain the concept of clustering in machine learning",</instruction>
37 <instruction>: "Explain the concept of direct preference optimization in machine learning",</instruction>
38 <instruction>: "What is the difference between a model and a predictive maintenance system?",</instruction>
39 <instruction>: "Explain the concept of data poisoning in machine learning",</instruction>
40 <instruction>: "Describe the concept of drug discovery in machine learning",</instruction>
41 <instruction>: "Explain the concept of contrastive language Image pre-training in machine learning",</instruction>
42 <instruction>: "Describe the concept of collaborative filtering in machine learning",</instruction>
43 <instruction>: "Describe the concept of Generative Adversarial Networks in machine learning",</instruction>
44 <instruction>: "Describe the concept of quantization in machine learning",</instruction>
45 <instruction>: "Describe the concept of thinking step by step in machine learning",</instruction>
46 <instruction>: "Explain the concept of supervised fine tuning in machine learning",</instruction>
```

aws

train.jsonl

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved. Amazon Confidential and Trademark.

You only need a
single jsonl file.



Preparing New Dataset

The screenshot shows a web application titled "League of LLM Data Set Generator". On the left is a dark sidebar with navigation links: "PR for Amazon" (selected), "Generate app", "Home", "Apps", "Snapshots", "Playlists", "Images", "PartyRock Guide", "What's new", "Roadmap", and "Backstage". The main content area has a green header bar with "League of LLM Data Set Generator" and icons for "Snapshot", "Remix", "Share", and a menu. Below the header is an "Introduction" section with text about generating training data for fine-tuning large language models. A "Topic Input" section contains a text input field with placeholder text: "→ Enter the topic or domain you want to generate training data for (e.g., 'Ancient Roman History', 'Python Programming', 'Climate Science')". A "Generated Questions" section at the bottom has a message: "Fill in Topic Input to get started!" and a button: "Fill in widgets to play".

<https://tinyurl.com/aws-ail-dataset-generator>



Automatic Training Synthetic Dataset Generator

PR for Amazon <

League of LLM Data Set Generator

[i](#) [Snapshot](#) [Remix](#) [Share](#) [≡](#)

[Generate app](#)

[Home](#)

[Apps](#)

[Snapshots](#)

[Playlists](#)

[Images](#)

[PartyRock Guide](#)

[What's new](#)

[Roadmap](#)

[Backstage](#)

Introduction

This tool helps you generate training data for fine-tuning large language models. Simply enter your topic or domain, and the system will generate 50 relevant questions and answers, formatted in JSONL format ready for fine-tuning.

Topic Input

→ Enter the topic or domain you want to generate training data for (e.g., 'Ancient Roman History', 'Python Programming', 'Climate Science')

1. Type "Topic" here

Generated Questions

Fill in Topic Input to get started!

Fill in widgets to play



Automatic Training Synthetic Dataset Generator

Generated Questions

- 1. What is the best time of year to visit New York City to avoid peak tourist crowds while still enjoying favorable weather?
- 2. How much should tourists budget per day for a mid-range vacation in Manhattan, including accommodation, food, and activities?
- 3. Which observation deck offers the best value for money: Empire State Building, Top of the Rock, or One World Observatory?
- 4. What are the lesser-known neighborhoods in Brooklyn that offer authentic local experiences?
- 5. How does the New York CityPASS compare to the New York Pass in terms of value and included attractions?
- 6. What are the most efficient transportation routes from JFK, LaGuardia, and Newark airports to Manhattan?
- 7. Which Broadway shows typically offer rush tickets or lottery systems for discounted seats?

Briefly inspect the generated questions



Generated Answers

```
{"instruction": "What is the best time of year to visit New York City to avoid peak tourist crowds while still enjoying favorable weather?", "context": "", "response": "Late September to early November offers mild temperatures and smaller crowds. Tourist numbers drop after Labor Day, while weather remains pleasant with temperatures around 60-70°F (15-21°C). Fall foliage adds beauty, especially in Central Park, and hotel rates begin to decrease from summer peaks."} {"instruction": "How much should tourists budget per day for a mid-range vacation in Manhattan, including accommodation, food, and activities?", "context": "", "response": "Budget approximately $250-350 per day for a mid-range Manhattan experience. This includes $150-200 for a decent hotel room, $50-75 for meals, $30-40 for attractions, and $20-30 for transportation. Prices can vary significantly based on season and specific neighborhood."} {"instruction": "Which observation deck offers the best value for money: Empire State Building, Top of the Rock, or One World Observatory?", "context": "", "response": "Top of the Rock offers the best value, providing panoramic views of both Central Park and the Empire State Building. Lines are typically shorter than ESB, and the three-level observation deck allows for less crowded viewing. Sunset visits offer particularly good value with stunning city light transitions."} {"instruction": "What are the lesser-known neighborhoods in Brooklyn that offer authentic local experiences?", "context": "", "response": "Explore Bushwick for street art and creative spaces, Red Hook for waterfront views and artisanal food, and Carroll Gardens for Italian-American culture. Greenpoint offers Polish heritage and trendy cafes, while Sunset Park features authentic Asian and Latin American communities."} {"instruction": "How does the New York CityPASS compare to the New York Pass in terms of value and included attractions?", "context": "", "response": "CityPASS is better for first-time visitors focusing on major attractions, offering 6 sites over 9 days at a lower price. New York Pass provides access to 100+ attractions but requires intensive sightseeing to maximize value. CityPASS averages $45 per attraction while New York Pass needs 2-3 daily visits to be worth it."} {"instruction": "What are the most efficient transportation routes from JFK, LaGuardia, and Newark airports to Manhattan?", "context": "", "response": "From JFK, take AirTrain to Jamaica Station, then LIRR to Penn Station (fastest) or subway. From LaGuardia, take Q70-SBS to subway connections. From Newark, take AirTrain to NJ Transit train to Penn Station. During rush hour, consider fixed-rate taxis to avoid traffic delays."}
```



Click on Copy Icon to copy the data



Briefly inspect the generated question-answer pairs.





© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved. Amazon Confidential and Trademark.

Section 5 - Training Data Generation

30



Final Round

The top 5 participants with the highest model win rates will move on to the Final Round, where they will compete using their own models

Dataset Preparation

Best Practices

Best Practices

- Fine tuning is a process that requires significant resources and cost, so we need to optimize it
- Find information that the model doesn't know yet or cannot answer before fine tuning to maximize win rate improvement
- Validate the dataset by testing it with baseline model to check each pair from the dataset
- If the model can answer the prompt correctly without being fine-tuned, then that pair should be removed from the dataset
- Collect enough dataset before submit for training (>11 KB is recommended) to avoid error.

Dataset Preparation with AWS Partyrock

Simple AWS AI League Dataset Generator

Intro

This playground will help you generate training data for fine-tuning large language models for the AWS AI League.

(An advanced version of the dataset generator can be found here: <https://partyrock.aws/u/pd7/pY9GNGEg2/Advanced-AWS-AI-League-Dataset-Generator>)

Topic to generate
→ Enter text relevant to your AI League competition category (i.e machine learning)

Number of data rows to generate
50

Questions

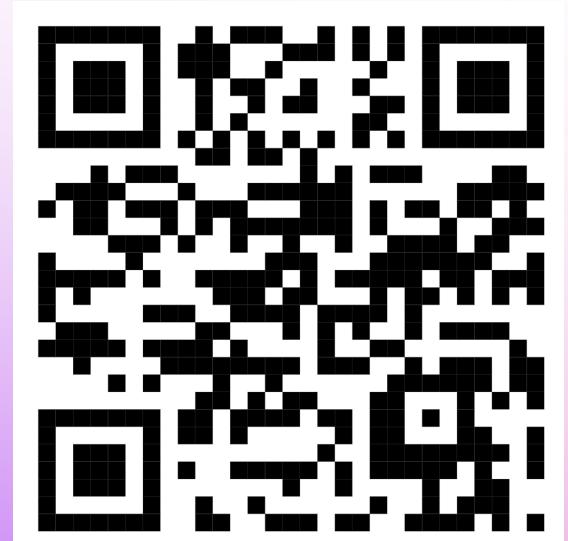
Fill in Topic to generate to get started!

Dataset

Fill in Topic to generate to get started!

Fill in widgets to play

- Quick generation of 50 pair prompt and response in jsonl for specified topic
- Very quick and easy to use! Just type the topic and run

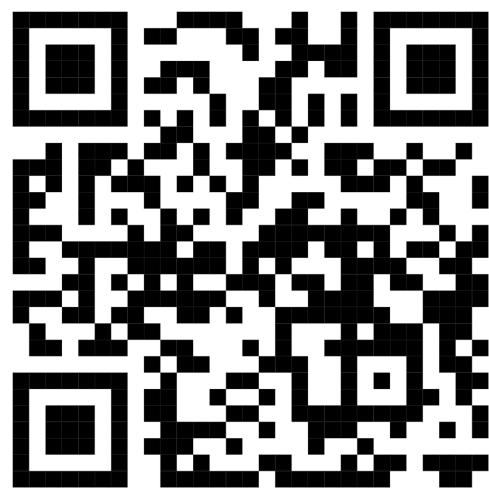


Dataset Preparation with Amazon Q CLI

AWS



Q CLI



- Can use Q Dev pro subscription or AWS BuilderID
- Longer time to generate dataset, but potentially higher quality via several prompting
- Create AWS BuilderID :
https://docs.aws.amazon.com/signin/latest/userguide/create-aws_builder_id.html
- Install Q CLI:
<https://docs.aws.amazon.com/amazonq/latest/qdevloperv-ug/command-line-installing.html>

Demo

Dataset Preparation

Next Steps!

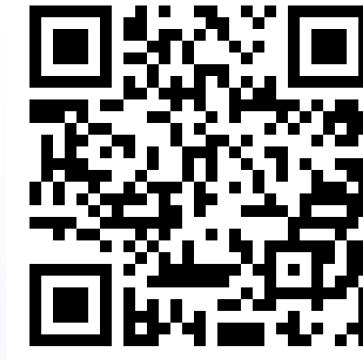
1. Prepare your dataset to bring to the event (using PartyRock or other tools such as Jakarta tourism website)
2. Reminder on important information that participants need to prepare on Summit event
3. Learn using these following resources:



AWS AI League Blog



Tips on AI League



SageMaker
Hyperparameter Docs



PartyRock

Q&A Time!

Drop your questions in the chat or raise your hand

Thank you!

