

AIM 206 - AWS AI League

AWS Summit Los Angeles

Mike Glover

(He/Him)

Senior Solutions Architect

Ash Naik

(She/Her)

Solutions Architect

Parth Desai

(He/him)

Startup Solutions Architect



Agenda

1. AWS AI League

Introduction to AWS AI League

2. Model Customization

Introduction to Model Customization

Hands-on: Model fine tuning

Hands-on: Model submission

3. Data Preparation

Introduction to **PARTYROCK**

Hands-on: Prepare your own synthetic dataset

4. Evaluation

Introduction to Model Evaluation

Next steps with improving model

5. Compete



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



AWS AI League

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



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

AWS AI League

- ★ Identify new Gen AI use cases

- ★ Build essential Gen AI capabilities using cutting edge AI tools

- ★ Foster friendly competition to drive innovation



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

Prizes



The background image shows the Las Vegas Strip at night, with the Venetian and Stratosphere towers prominently visible. Overlaid on the center is a large white rectangular box containing the AWS re:Invent logo and event details.

AWS re:Invent
DECEMBER 1–5, 2025 | LAS VEGAS, NV

Win re:Invent travel package to compete
in the **AWS AI League Final**
+ Top 25 will receive \$75 AWS Credits



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

AWS AI League overview



2 HRS

Hands-on Workshop

Model Customization
Dataset Preparation
Evaluation

8:30 - 10:30 PT



< 10 HRS

Virtual Competition

Compares model performance for a specific domain using relevant data

10:30 - 4:00 PM PT



30-45 MINS

Gen AI Gameshow

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

Preview Experience at AI Hub



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



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



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

AI League Challenges



PROMPT SAGE



TUNE WHIZ



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

The screenshot shows the AWS AI League interface for the New York Summit 2025. The top navigation bar includes 'AWS AI League' and 'AWS AI League' tabs, and a 'Atlas2025' button. On the left, a sidebar provides links for 'Artifact Management', 'Your Trainer Profile', 'Leaderboard Submission' (which is highlighted in blue), and 'Validate Training Data'. The main content area displays the 'leaderboard' for the 'AWS AI League - New York Summit 2025'. It shows a total of 352 trainers. A message at the top states: 'You have no pending artifact submission. To evaluate a new artifact to the leaderboard, submit your new model again.' Below this is a search bar labeled 'Find resources' and a pagination control showing page 1 of 18. The main table lists the top 5 trainers:

Rank	Trainer Alias	Best Artifact ID	Best Win Rate	Gap to First	Last Submitted Win Rate	No. of Wins	No. of Total Runs	Total Submissions
1	Naomi	model983	89%	0%	85%	42	49	41
2	Wilbur	distilled023	87%	0%	34%	17	49	20
3	Natasya	aiTrain100	85%	4%	85%	42	49	4
4	Vincent	dmind379	81%	8%	81%	40	49	25
5	Agnes	sml238	77%	12%	77%	38	49	56



TUNE WHIZ CHALLENGE

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

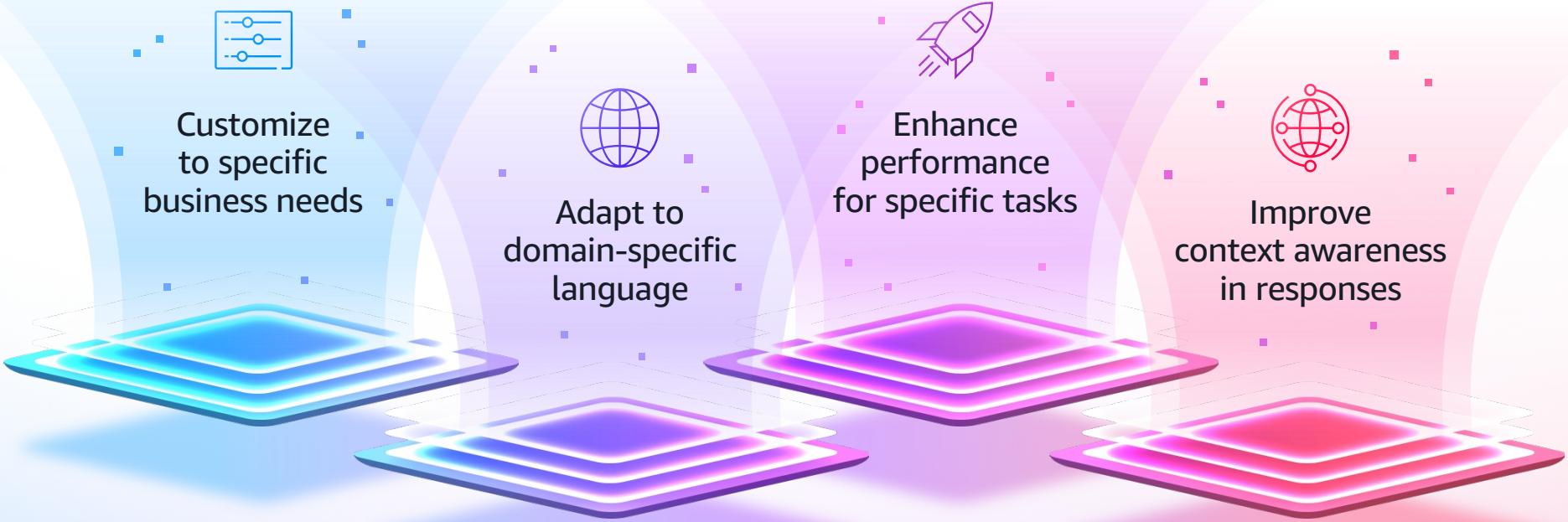
CHALLENGE

- 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.



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

Why Customizing Foundational Model (FM) is important?



Contest Theme

Los Angeles Tourist Information Virtual Assistant – LA.Go



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

LA Tourist Information Virtual Assistant - Your Personal LA Guide

CHALLENGE

Los Angeles 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 "**Los Angeles voice**" that resonates with visitors, 'Eh'.

SOLUTION

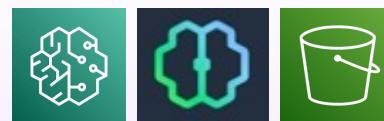
Host an AWS AI League event to develop "**LA.go**" - a custom-tuned virtual tourism assistant that captures authentic Los Angeles voice. This cost-effective model will be trained on local content providing visitors with accurate, culturally-informed guidance.

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 Los Angeles experience.



KEY SERVICE(S)



Amazon SageMaker, SageMaker Unified Studio & Amazon S3



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

What the Tourism office really wants

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



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

Example

What's the best way mode of transportation to see the top sights in Los Angeles? And I don't want to be caught as a tourist.



Base Model

- The tourist is informed that there are 7 different modes of transportation, public transit, streetcar, bike, walking, hop-on hop off bus tour, taxi, tour services.
- The tourist is informed of multiple different sights.
- The tourist is tipped to use Google.

Fine Tuned Response - Potential

What's the best way mode of transportation to see the top sights in LA? And I don't want to be caught as a tourist.



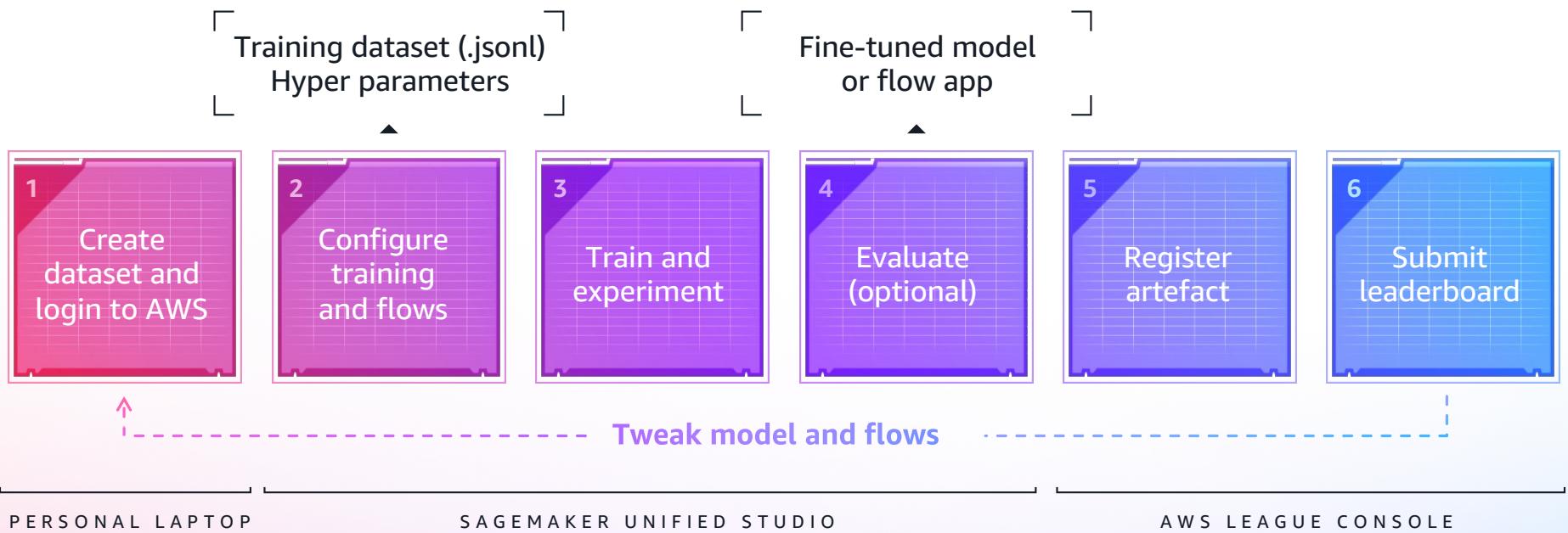
For the top sights in Los Angeles, you've got to check out the **Hollywood Sign and Walk of Fame**.

Visit Griffith Observatory for **amazing city views** and stargazing opportunities. From there, you can also hike in Griffith Park or check out the Hollywood Sign up close.

For **culture**, explore the Getty Center's amazing art collection and gardens, or walk through the trendy Arts District downtown with its murals, galleries, and hip restaurants. Don't miss iconic spots like Santa Monica Pier, **Universal Studios**, or catching a show at the Hollywood Bowl during summer.

For a **local experience**, grab some tacos from a food truck, browse the Grand Central Market downtown, or explore neighborhoods like Silver Lake or Los Feliz where **Angelenos hang out**.

AWS AI League flow diagram



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

How to prepare new tuning datasets



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

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.

```
Users/vincenzoh/Downloads/project-lol/instruction-tuning/jsonl_files/random50.jsonl
10 {"instruction": "What is the difference between a model and a dependency parser?", "context": "", "response": "A model, in the context of Machine Learning, refers to a system that can learn from and make predictions on data. A dependency parser, on the other hand, is a type of natural language processing (NLP) model that identifies the grammatical dependencies between words in a sentence."}
11 {"instruction": "Explain the concept of jailbreaking in machine learning.", "context": "", "response": "Jailbreaking in machine learning appears to be a misnomer or a metaphorical term. In machine learning, alignment to human preferences is a fundamental goal. If a model's predictions align well with human preferences, it is considered 'aligned'. If they do not, it is considered 'misaligned'."}
12 {"instruction": "Describe the concept of alignment to human preferences in machine learning.", "context": "", "response": "Alignment to human preferences is a key goal in machine learning. It means that the model's predictions should be consistent with what humans expect or prefer. This is often achieved by training the model on data that reflects human preferences."}
13 {"instruction": "What is the difference between a model and a reinforcement learning system?", "context": "", "response": "A model is a representation of the world or a process, while a reinforcement learning system is a specific type of model that learns by interacting with an environment to maximize a reward signal."}
14 {"instruction": "Explain the concept of low-rank induction in machine learning.", "context": "", "response": "Low-rank induction is a technique used in machine learning to find a low-rank approximation of a matrix. This is often done to reduce the computational complexity of the model or to capture the most important features in the data."}
15 {"instruction": "What is the difference between a model and a hypothesis?", "context": "", "response": "A hypothesis is a statement or claim that is being tested or evaluated. A model, on the other hand, is a representation of the world or a process that is used to make predictions or decisions based on data."}
16 {"instruction": "Explain the concept of clustering in machine learning.", "context": "", "response": "Clustering is a type of unsupervised machine learning algorithm that groups data points into clusters based on their similarity. The goal is to find inherent groupings in the data without prior knowledge of the categories."}
17 {"instruction": "Describe the concept of explainability in machine learning.", "context": "", "response": "Ah, an excellent topic! Explainability, or interpretability, refers to the ability to understand how a machine learning model arrived at its conclusions. It's important for building trust in AI systems and for ensuring that the model's decisions are fair and unbiased."}
18 {"instruction": "Explain the concept of backtracking in machine learning.", "context": "", "response": "Backtracking is a search strategy used in machine learning to explore different paths or hypotheses. It involves starting with a hypothesis and then iteratively refining it by adding or removing features until a good fit is found."}
19 {"instruction": "Describe the concept of precision and recall in machine learning.", "context": "", "response": "Precision and recall are two important metrics used to evaluate the performance of classification models. Precision measures the proportion of true positives among all predicted positives, while recall measures the proportion of true positives among all actual positives."}
20 {"instruction": "Describe the concept of alignment to human preferences in machine learning.", "context": "", "response": "In machine learning, alignment to human preferences is a key goal. It means that the model's predictions should be consistent with what humans expect or prefer. This is often achieved by training the model on data that reflects human preferences."}
21 {"instruction": "Describe the concept of latent diffusion models in machine learning.", "context": "", "response": "Latent diffusion models are a type of generative model used in machine learning. They work by first generating a random noise vector and then applying a series of learned transformations to it to produce a realistic image or other output."}
22 {"instruction": "Describe the concept of dialogue systems in machine learning.", "context": "", "response": "Ah, an excellent topic! Dialogue systems, also known as conversational AI, are computer programs designed to engage in natural language conversations with humans. They use machine learning techniques to understand and generate human-like responses."}
23 {"instruction": "Explain the concept of knowledge distillation in machine learning.", "context": "", "response": "Knowledge distillation is a technique in machine learning where a complex, large model (the "teacher") is trained to teach a smaller, simpler model (the "student"). The student model then learns to produce similar results to the teacher but with fewer resources."}
24 {"instruction": "Describe the concept of merging large language models (LLMs) in the context of merging large language models (LLMs).", "context": "", "response": "Great question! As you mentioned, merging large language models (LLMs) is a way to combine the strengths of multiple LLMs into a single, more powerful model. This can be achieved through various methods like averaging, weighted averaging, or more advanced techniques like knowledge distillation."}
25 {"instruction": "What is the difference between a model and a generative system?", "context": "", "response": "Great question! A generative system is a type of model that can generate new data samples that are similar to the training data. It does this by learning the underlying probability distribution of the data."}
26 {"instruction": "What is the difference between a convolutional neural network and a recurrent neural network?", "context": "", "response": "Ah, an excellent topic! Convolutional neural networks (CNNs) are typically used for processing spatial data like images, while recurrent neural networks (RNNs) are used for sequential data like text or speech. CNNs use local receptive fields and shared weights across the entire input, while RNNs have a recurrent connection that processes the sequence step-by-step."}
27 {"instruction": "What is the difference between a support vector machine and a decision tree?", "context": "", "response": "Great question! As a professor who specializes in machine learning, I can tell you that support vector machines (SVMs) and decision trees are two very different types of models. SVMs are based on linear or non-linear decision boundaries, while decision trees are hierarchical structures that split data into smaller and smaller subsets based on feature values."}
28 {"instruction": "What is the difference between a model and a framework?", "context": "", "response": "Great question! As a professor who specializes in machine learning, I can tell you that a model is a specific implementation of a learning algorithm, while a framework is a general-purpose system that provides the infrastructure for building machine learning models."}
29 {"instruction": "Describe the concept of selective state spaces in machine learning.", "context": "", "response": "Selective state spaces are a concept in machine learning that refers to the idea of only considering certain states or actions in a problem space. This can be useful for reducing the search space and improving efficiency."}
30 {"instruction": "Explain the concept of reward models in machine learning.", "context": "", "response": "Reward models are a type of machine learning model that are used to predict the value or reward of different actions or states. They are often used in reinforcement learning to guide the learning process."}
31 {"instruction": "Describe the concept of adversarial training in machine learning.", "context": "", "response": "Adversarial training is a technique used in machine learning to improve the robustness of a model. It involves training the model on a dataset that includes both normal data and adversarial examples (data that has been modified to intentionally fool the model). The model is then trained to correctly classify both types of data."}
32 {"instruction": "Describe the concept of precision and recall in machine learning.", "context": "", "response": "Precision and recall are two important metrics used to evaluate the performance of classification models. Precision measures the proportion of true positives among all predicted positives, while recall measures the proportion of true positives among all actual positives."}
33 {"instruction": "Describe the concept of multi-head attention in machine learning.", "context": "", "response": "Multi-head attention is a mechanism used in transformer models to allow the model to attend to different parts of the input sequence simultaneously. It consists of multiple parallel attention heads, each of which performs a weighted sum of the input features based on their relevance to the query."}
34 {"instruction": "Explain the concept of reward models in machine learning.", "context": "", "response": "Reward models are a crucial component of reinforcement learning. They provide a way to measure the quality of an action taken by the agent, which is then used to guide the learning process."}
35 {"instruction": "Explain the concept of tokenization in machine learning.", "context": "", "response": "Tokenization is a preprocessing step in natural language processing (NLP) that involves breaking down text into smaller, meaningful units called tokens. These tokens can then be processed by a machine learning model to extract meaning."}
36 {"instruction": "Explain the concept of clustering in machine learning.", "context": "", "response": "Ah, a most excellent topic! Clustering is a fundamental concept in machine learning that refers to the process of grouping data points into clusters based on their similarity. It is often used for tasks like market segmentation or document clustering."}
37 {"instruction": "Explain the concept of the difference between a machine learning model and a generative model.", "context": "", "response": "The main difference between a machine learning model and a generative model is that a machine learning model is trained to predict the next data point in a sequence, while a generative model is trained to generate new data points that are similar to the training data."}
38 {"instruction": "What is the difference between a generative model and a discriminative model?", "context": "", "response": "Ah, an excellent question! A generative model is a type of model that generates new data samples, while a discriminative model is a type of model that classifies data into different categories based on its features."}
39 {"instruction": "Explain the concept of data poisoning in machine learning.", "context": "", "response": "Data poisoning is a type of adversarial attack in machine learning where an attacker injects malicious data into the training set to corrupt the model's training process."}
40 {"instruction": "Describe the concept of drug discovery in machine learning.", "context": "", "response": "Drug discovery in machine learning is a process where machine learning models are used to predict the properties of potential drugs. This can involve tasks like predicting the binding affinity of a drug candidate or identifying new drug targets."}
41 {"instruction": "Explain the concept of contrastive language image pre-training in machine learning.", "context": "", "response": "Contrastive language image pre-training is a type of machine learning task where the model is trained to distinguish between pairs of images and their corresponding captions. This is often used to improve the performance of image captioning and other NLP tasks."}
42 {"instruction": "Describe the concept of collaborative filtering in machine learning.", "context": "", "response": "Ah, an excellent topic! Collaborative filtering is a recommendation system that uses the ratings or preferences of users to predict what they might like in the future."}
43 {"instruction": "Describe the concept of Generative Adversarial Networks in machine learning.", "context": "", "response": "Generative Adversarial Networks (GANs) are a type of machine learning model that consists of two neural networks: a generator and a discriminator. The generator creates new data samples, while the discriminator tries to distinguish between real data and generated data."}
44 {"instruction": "Describe the concept of quantization in machine learning.", "context": "", "response": "Ah, an excellent topic! Quantization is a fundamental technique used in machine learning to reduce the number of bits required to represent data, which can lead to faster computation and lower memory usage."}
45 {"instruction": "Describe the concept of thinking step by step in machine learning.", "context": "", "response": "Thinking step by step is an excellent question! Thinking step by step is a common approach in machine learning, especially for tasks like reinforcement learning, where the model needs to consider the consequences of its actions over time."}
46 {"instruction": "Explain the concept of supervised fine tuning in machine learning.", "context": "", "response": "Supervised fine-tuning is a technique used in machine learning to improve the performance of a pre-trained model on a specific task. It involves taking a pre-trained model and fine-tuning its parameters on a labeled dataset for the specific task at hand."}
```

You only need a
single jsonl file. 😊



train.jsonl

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

STAGE 1

Dataset preparation with PartyRock

The screenshot shows the 'League of LLM Data Set Generator' interface. On the left is a dark sidebar with navigation links: 'Generate app', 'Home', 'Apps', 'Snapshots', 'Playlists', 'Images', 'PartyRock Guide' (with 'What's new' and 'Roadmap' subsections), and 'Backstage'. The main area has a green header bar with icons for 'Snapshot', 'Remix', 'Share', and a menu. Below the header, the 'Introduction' section explains the tool's purpose: generating training data for fine-tuning large language models. The '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')'. At the bottom, the 'Generated Questions' section displays a message: 'Fill in `Topic Input` to get started!' next to a play button icon and a 'Fill in widgets to play' button.



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

Automatic Training Synthetic Dataset Generator

The screenshot shows a user interface for generating training data. At the top, there's a green header bar with the text "Generated Questions". Below it is a list of 7 questions about New York City tourism:

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?

To the right of the questions, a red callout box with a red arrow points to the text: "Briefly inspect the generated questions".

At the bottom, there's a green header bar with the text "Generated Answers". Below it is a large block of JSON-formatted text representing the responses to the questions above. The responses include detailed explanations for each question, such as the best time to visit NYC, budgeting tips, and transportation routes.

To the right of the answers, another red callout box with a red arrow points to the "Copy" icon (a small square with a double-lined border) in the top right corner of the answer area, with the text: "Click On Copy Icon to copy the data".

A third red callout box with a red arrow points to the text: "Briefly inspect the generated question-answer pairs.", located to the right of the JSON text.

```
{"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."}
```

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

The diagram illustrates a four-step process for fine-tuning a model using the AWS SageMaker JumpStart interface. Step 1 shows the 'Providers' section where the 'Meta' provider is selected. Step 2 shows the 'Train Model' configuration page with the 'Data Source' section highlighted. Step 3 shows the 'Hyper-parameters' section being customized. Step 4 shows the 'Train' button, indicating the no-code, one-click tuning capability.



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

Training Data Validation

AWS AI League AWS AI League

AWS AI League

Models Management

Your Trainer Profile

Leaderboard Submission

Validate Training Data

1. Click "Validate Training Data"

2. Paste generated data from PartyRock or other sources into JSONL editor

Paste JSONL Data

Paste your training data in JSONL format for validation.

```
32 es, AI-powered tutoring, and integration with industry-standard tools. Additionally, platforms could expa  
33 lts of their decisions. This approach helps students understand the practical implications of AI/ML theor  
34 h learning objectives and accommodating diverse learning styles. These challenges could be addressed thro  
35 ugh kinesthetic learners. Additionally, platforms might include collaborative and competitive modes, all  
36 real time, and offer personalized feedback. This approach may lead to more efficient learning, increased  
37 motivation, complement traditional coursework, and facilitate project-based learning. Furthermore, platforms  
38 can reward ethical choices and provide feedback on the impact of AI on society. By addressing these challenges, we  
39 can help students develop a deeper understanding of AI and its applications in various fields.  
40 lowing and decrease it as the distance increases. Return the calculated reward value."  
41 e this custom action space in your DeepRacer model."  
42 erence. Update the target waypoint when reached. Return steering and speed commands."  
43 olean indicating whether the agent is on or off the track."  
44 . Implement a penalty for sudden speed changes. Combine this with other reward factors like track positio  
45 n of future positions. This can be used for collision prediction or advanced path planning."  
46 ity. Return modified steering and speed commands to navigate around obstacles safely."  
47 s. Combine this with other reward factors like track position and speed to create a balanced training ob  
48 target speeds. Gradually adjust current speed towards the target. Return the adjusted speed command."  
49  
50 {"instruction5": "Implement a function to dynamically adjust the DeepRacer's speed based on upcoming trac
```

3. Validate JSONL

Upload jsonl file

Valid training data format!

Validate JSONL

Save Training Data File

4. Save Training Data file (i.e. train.jsonl) to local disk

5. In the event of a validation error, fix the error and validate again.



50 {"instruction5": "Implement a function to dynamically adjust the DeepRacer's speed based on upcoming trac

Upload jsonl file

⚠ Error on line 50: Missing required keys: instruction

Validate JSONL Save Training Data File

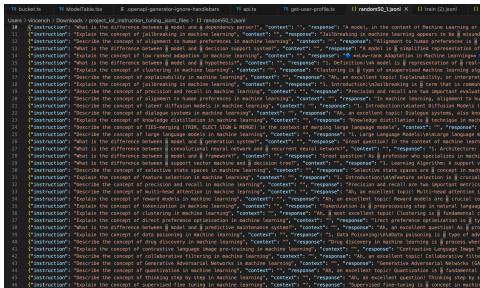


© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved. Amazon Confide

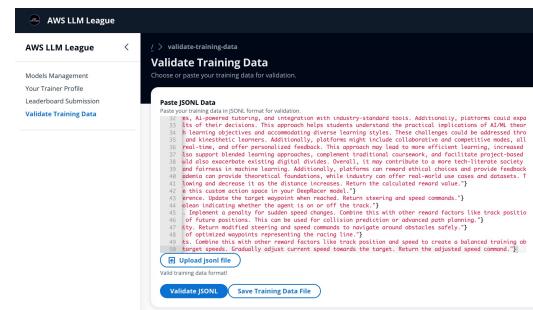
SECTION 3 - TRAINING DATA GENERATION

Upload into S3

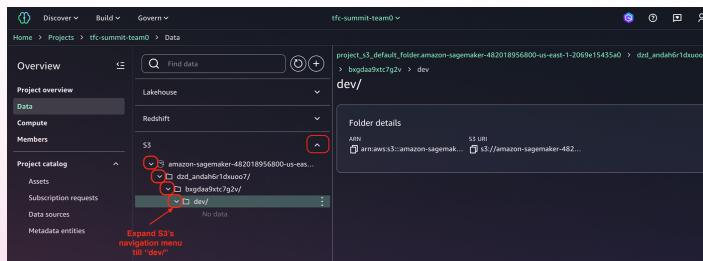
Note: Step by step guide in Workshop studio instructions



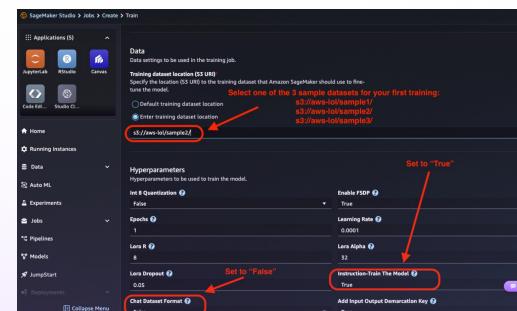
Step1. Prepare dataset and name it locally “train.jsonl”



Step2. Validate dataset at aws-lm-league.com



Step3. Upload dataset into SageMaker Unified Studio's Project



Step4. Create a new training job with your S3 URI as the input dataset

Fine-tuning model with JumpStart

The diagram illustrates a four-step process for fine-tuning a model using AWS SageMaker JumpStart:

- 1 Select model from catalogue**: Shows the "Providers" section of the SageMaker Public Hub interface. A callout box highlights the "Meta" provider.
- 2 Provide training dataset**: Shows the "Train Model" configuration page for the Llama-2-7b model. A callout box highlights the "Data Source" section where the "Training data set" is specified as `s3://jumpstart-cache-prod-us-west-2/training-datasets/sec_amazon/`.
- 3 Customize hyperparameters**: Shows the "Hyper-parameters" section of the configuration page, highlighted by a callout box.
- 4 No-code one-click tuning**: Shows the "Train" button at the bottom left of the configuration page, highlighted by a callout box.



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

How to submit a model to leaderboard



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

Registering and Submitting Models

The screenshot shows the 'Model Registration' step of the AWS AI League interface. It includes fields for 'Model Name' (Noor Reinvent Model), 'Training Job Name' (Noor-Model2), and 'Select SageMaker' (SageMaker). A red arrow points to the 'Model Name' field. Another red arrow points to the 'Select SageMaker' dropdown, which is currently set to 'SageMaker'. The sidebar on the left lists 'Models Management', 'Your Trainer Profile', 'Leaderboard Submission', and 'Validate Training Data'.

Step 1 – Register Model

The screenshot shows the 'Leaderboard' page for the 'Reinvent24 Large Language Models League'. It features a table with columns for Rank, Trainer Alias, Best Model ID, Best Win Rate, Gap to First, Last Submitted Win Rate, No. of Wins, No. of Total Runs, and Total Submissions. A red box highlights the '+ Submit your model' button at the top right of the page. The table data is as follows:

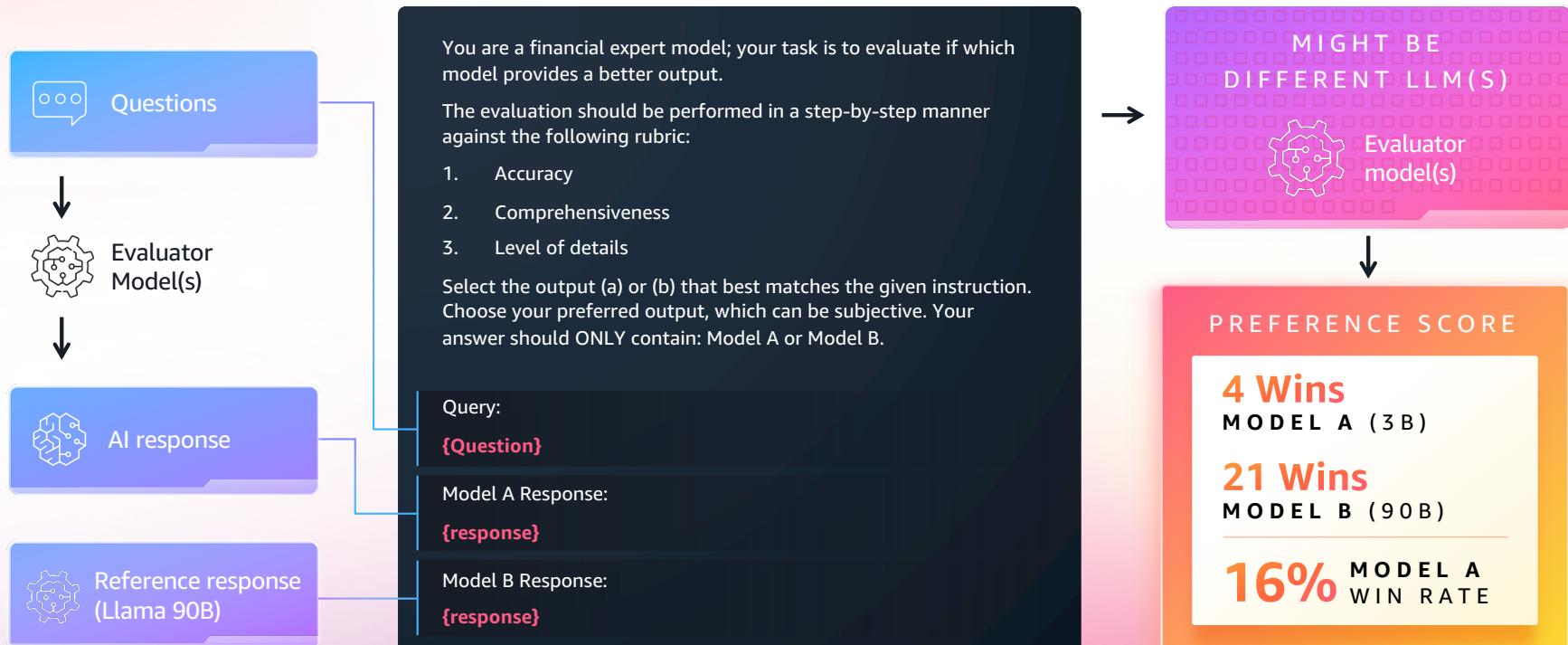
Rank	Trainer Alias	Best Model ID	Best Win Rate	Gap to First	Last Submitted Win Rate	No. of Wins	No. of Total Runs	Total Submissions
1	AWS-Vincent	622dbba2-c285-4676-89f5-530e5c0c6f18	40%	0%	16%	8	49	4

Step 2 – Submit Model



AI Critique

LLM-based response scoring



Hands On Time



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

AWS AI League – AWS Summit LA

Workshop Links

URL: <https://s12d.com/ai-league-la>



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

Next steps

Now (This Session and until 12:30 PM):

1. Experience tuning, evaluating & submitting first model to AWS AI League competition
2. Subsequently train with different set of data & hyper parameters
3. Submit your new trained models to leaderboard to achieve better rankings

12:30 PM PT:

1. Check Leaderboard if you are on the Top 25

12:30 PM PT:

1. Top 25 claim their prize at the AWS AI League booth at the AI Hub

After the Session:

1. Preview experience of the Gen AI Gameshow at the AWS AI League booth at the AI Hub
2. Share your experience on Social Media
3. Explore how we can bring this experience to your organization
4. Compete at AWS AI League at re:Invent 25



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

Thank you!

Mike Glover

mrglover@amazon.com

Ash Naik

ashjnaik@amazon.com

Parth Desai

parthdes@amazon.com

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

