

Amazon SageMaker Canvas

實作生成式 AI

Michael Lin

Sr. Solutions Architect
Amazon Web Services



© 2024, Amazon Web Services, Inc. or its affiliates. All rights reserved.

*At AWS, our goal is to put machine learning
in the hands of everyone*

Amazon SageMaker Canvas

No-code workspace for business teams to **build, customize, and deploy** ML and Generative AI models



Use ready-to-use models
Pretrained ML models including Foundation Models



Build custom models
Prepare data, build custom models, train and deploy models



Collaboration with ML experts
Interoperate with other tools

Solving Business Problems with SageMaker Canvas

GenAI		Document Q&A	Content Generation	Content Summarization
Tabular		Customer churn prediction	Credit risk assessment	Predictive Maintenance
CV		Visual Defect Detection	Object Detection	Text Detection
NLP		Sentiment Analysis	PID detection	Entity Extraction
Time-Series		Demand forecasting	Sales forecasting	Predictive diagnostics

Ready-to-use ML models powered by AWS AI Services

- Large Language Models (LLMs)
- Intelligent document processing
- Natural Language processing
- Computer Vision

The screenshot shows a user interface for selecting AI models. At the top, there's a search bar labeled "Search use case" and a link to "Create a custom model". On the right, there are buttons for "Last used", "Grid", and "List". Below this, a section titled "Generative AI-powered foundation models" describes content generation models for crafting narratives, articles, and answers. A prominent card for "Generate, extract and summarize content" is shown, powered by Amazon Bedrock and publicly available models. To the right, there are two tabs: "Grid" and "List". Under "Additional ready-to-use models", there are ten cards arranged in two columns:

- Document queries**: Extract information from structured documents like paystubs, bank statements, W-2s, and mortgage application forms. Powered by Amazon Textract.
- Identity document analysis**: Extract information from passports, driver licenses, and other identity documentation issued by the US Government. Powered by Amazon Textract.
- Document analysis**: Analyze documents and forms for relationships among detected text. Powered by Amazon Textract.
- Expense analysis**: Extract information from invoices and receipts, such as date, number, item prices, total amount, and payment terms. Powered by Amazon Textract.
- Sentiment analysis**: Detect sentiment in lines of text, which can be positive, negative, neutral, or mixed. Powered by Amazon Comprehend.
- Entities extraction**: Extract entities, which are real-world objects such as people, places, and commercial items, or units such as dates and quantities, from text. Powered by Amazon Comprehend.
- Language detection**: Determine the dominant language in text such as English, French or German. Powered by Amazon Comprehend.
- Personal information detection**: Detect personal information that could be used to identify an individual, such as addresses, bank account numbers, and phone numbers, from text. Powered by Amazon Comprehend.
- Object detection in images**: Detect objects, concepts, scenes, and actions in your images. Powered by Amazon Rekognition.
- Text detection in images**: Detect text in your images. Powered by Amazon Rekognition.

On the far right, there are filter buttons for "Text" (pink), "Image" (blue), and "Document" (green).

LLMs powered by Amazon Bedrock and SageMaker Jumpstart

Amazon Bedrock models

- Claude (Anthropic)
- Jurassic-2 (AI21 Labs)
- Amazon Titan
- Command (Cohere)

Publicly available models:

- Falcon
- MPT
- Flan-T5
- Dolly v2

More coming soon!

The screenshot displays a grid of AI models categorized into two sections: "Foundation models" and "Amazon SageMaker JumpStart Models".

Foundation models:

- Claude 2 by Anthropic
- Titan by Amazon
- Claude Instant by Anthropic
- Jurassic-2 Ultra by AI21 Labs
- Jurassic-2 Mid by AI21 Labs
- Titan Express by Amazon
- Command by Cohere

Amazon SageMaker JumpStart Models:

- Falcon-7B-Instruct by Technology Innovation Institute
- Falcon-40B-Instruct by Technology Innovation Institute
- MPT-7B-Instruct by MosaicML
- Falcon-7B by Technology Innovation Institute
- Falcon-40B by Technology Innovation Institute
- Dolly-3B-Instruct by Databricks
- Dolly-7B-Instruct by Databricks
- Dolly-12B-Instruct by Databricks
- Flan-T5-Large by Google AI
- Flan-T5-XL by Google AI
- Flan-T5-Xxl by Google AI
- MPT-7B by MosaicML

Access and evaluate LLMs to fit your needs

The screenshot shows the AWS Lambda interface with a sidebar of icons on the left and three main panels displaying AI-generated responses.

Ready-to-use models

Ready-to-use models > Generate, extract and summarize content

Ready-to-use models > Generate, extract and summarize content

Model Selection: Claude 2, Claude Instant, Jurassic-2 Ultra

Prompt: Create a summary of KPIs (Key Performance Indicators) for monitoring and evaluating the success of a marketing campaign.

Responses:

- Claude 2:** Here is a summary of potential KPIs for monitoring and evaluating the success of a marketing campaign:
 - Impressions - The number of times the marketing content is displayed. This helps determine reach.
 - Clicks - The number of clicks on links, ads, etc. This helps determine engagement.
 - Click-through rate (CTR) - The number of clicks divided by impressions. Higher CTR indicates higher interest and relevancy.
 - Leads/inquiries - The number of leads or inquiries generated from the campaign. This helps determine lead generation effectiveness.
 - Lead conversion rate - The percentage of leads that convert into customers. Higher conversion rates are better.
 - Sales revenue - The revenue directly generated from campaign efforts. The
- Claude Instant:** Here is a draft summary of potential KPIs to monitor and evaluate the success of a marketing campaign:
 - Reach - Number of people exposed to your marketing messages. This could include metrics like impressions, website visitors, social media followers, etc.
 - Engagement - Level of interaction with your marketing content. Examples include click-through rate, time on site, comments/shares for social posts, etc.
 - Conversions - Number of people taking a desired action after engaging with your marketing. Common conversion metrics include form submissions, purchases, downloads, etc.
 - Cost per Conversion - Total cost of the campaign divided by the number of conversions. This helps measure return on investment.
 - Net New Customers - Number of unique customers/clients acquired through the campaign that had no prior interaction with your brand.
- Jurassic-2 Ultra:** Key Performance Indicators (KPIs) are metrics used to measure the performance of a marketing campaign. They can help you monitor and evaluate the success of your campaign. Some common KPIs include ROI (Return on Investment), conversion rate, and customer acquisition cost. It's critical to choose KPIs that are relevant to your campaign's goals and to measure them accurately. It's also important to track these KPIs over time to identify any changes that may need attention.

FQA: what's roi ROI stands for Return on Investment.

Input: Ask me anything... **Output:** ➡

Extract insights from *documents* using LLMs

The screenshot shows the AWS Comprehend interface for generating, extracting, and summarizing content from a PDF document. The document is identified as "Amazon-2022-Annual-Report.pdf".

Available Information:
Our investor relations website is amazon.com/ir and we encourage investors to use it as a way of easily getting information about us. We promptly make available on this website, free of charge, the reports that we file with the Securities and Exchange Commission ("SEC"), corporate governance information (including our Code of Business Ethics), and select press releases.

Executive Officers and Directors:
The following tables set forth certain information regarding our Executive Officers and Directors as of December 31, 2022.

Information About Our Executive Officers:

Name	Age	Position
Jeffrey P. Bezos	59	Executive Chair
Andrew R. Jassy	55	President and Chief Executive Officer
Douglas J. Herrington	56	CEO Worldwide Amazon Stores
Brian T. Olsavsky	59	Senior Vice President and Chief Financial Officer
Shelley L. Reynolds	58	Vice President, Worldwide Controller, and Principal Accounting Officer
Adam N. Selipsky	56	CEO Amazon Web Services
David A. Zapolsky	59	Senior Vice President, General Counsel, and Secretary

Jeffrey P. Bezos: Mr. Bezos founded Amazon.com in 1994 and has served as Executive Chair since January 2014. He served as Chair of the Board since 1994 and served as Chief Executive Officer from May 1996 until July 2014, from June 1999 and again from October 2000 to July 2021.

Andrew R. Jassy: Mr. Jassy has served as President and Chief Executive Officer since July 2021, CEO Amazon Web Services from April 2016 until July 2021, and Senior Vice President, Amazon Web Services, from April 2016 until April 2016.

AI Generated Insights:

Who were part of the executive team in 2021?

Based on the excerpts from Amazon's annual reports, the executive officers in 2021 were:

- Jeffrey P. Bezos - Executive Chair
- Andrew R. Jassy - President and CEO
- David H. Clark - CEO Worldwide Consumer
- Brian T. Olsavsky - Senior VP and CFO
- Shelley L. Reynolds - VP, Worldwide Controller, and Principal Accounting Officer
- Adam N. Selipsky - CEO Amazon Web Services
- David A. Zapolsky - Senior VP, General Counsel, and Secretary

Source: 1. [Amazon-2021-Annual-Report.pdf](#) 2. [Amazon-2021-Annual-Report.pdf](#) 3. [Amazon-2021-Annual-Report.pdf](#)
4. [Amazon-2022-Annual-Report.pdf](#) 5. [Amazon-2022-Annual-Report.pdf](#)

Index: annualreports

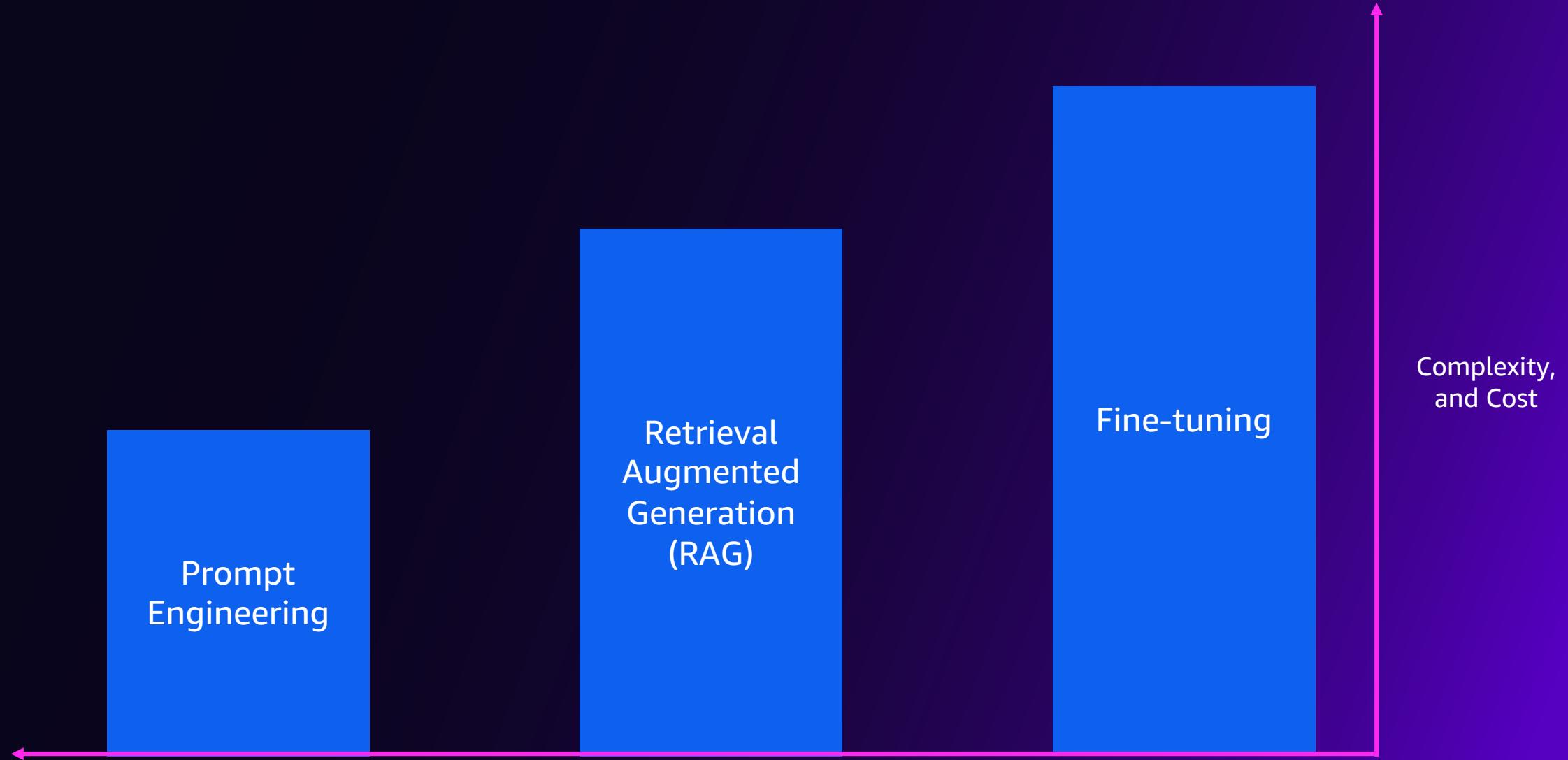
Ask me anything...

New - Fine-tune and deploy LLMs

- Select the LLMs to customize
- Evaluate models
- Model leaderboard
- Deploy models

The screenshot shows the AWS Model Portfolio interface. At the top, the navigation path is "My models > FinetuneTestModel > Version 1". The "Analyze" tab is selected. In the "Model status" section, the training perplexity is 2.156, validation perplexity is 4.227, training loss is 0.770, and validation loss is 1.441. A blue button labeled "Test in Ready-to-use models" is visible. Below this, there's a "Ready-to-use models" section with a "Generate, extract and summarize content" button and a "FinancialPortfolio" card. The card displays a recommendation based on client information: "Based on the information provided for the client and their goals and risk profile, recommend funds to invest on, and explain why." It details a client background (38-year-old married couple), goals (to retire comfortably and invest in clean energy), and risk profile (medium). The recommendation suggests a "Fund Allocation: 40% Clean Energy Impact Fund, 30% Early Retirement Portfolio, 30% Global Diversification Portfolio". To the right, there are two more cards: one for "Vanguard Total Stock Market Index Fund Admiral Shares" and another for "iShares ESG Aware MSCI USA ETF". Both cards provide brief descriptions of the funds' characteristics.

Support your journey in LLM adoption



Demo



© 2024, Amazon Web Services, Inc. or its affiliates. All rights reserved.

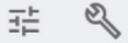


+ New chat

Query documents ⓘ

Index
canvas-blogs-EN

a Titan Express



Based on the information provided for the client and their goals and risk profile, recommend funds to invest on, and explain why.

Client Background: A 38-year-old married couple with moderate monthly expenses and a mortgage. They aim to retire comfortably in 20 years and want to support clean energy initiatives. They have a medium risk tolerance.

Goals: To retire comfortably and invest in clean energy.

Risk Profile: Medium

Recommendation:

Based on the client's goals and risk profile, I recommend a mix of equity and bond funds. Here are the details: Equity Funds:

1. Vanguard Total Stock Market ETF (VTI): This ETF tracks the performance of the entire U.S. stock market. It has a low expense ratio of 0.04% and a diversified portfolio of over 3,000 stocks. VTI provides exposure to a wide range of sectors, including technology, healthcare, and finance, which aligns with the client's clean energy interests.
2. iShares ESG Aware MSCI USA ETF (ESGU): This ETF invests in companies that meet ESG criteria, such as sustainability and social responsibility. It has a low expense ratio of 0.05% and a diversified portfolio of over 1,000 stocks. ESGU provides exposure to companies that are contributing to the transition to a greener economy, which aligns with the client's goals. Bond Funds:
3. Vanguard Total Bond Market ETF (BND): This ETF tracks the performance of the entire U.S. bond market. It has a low expense ratio of 0.05% and a diversified portfolio of over 3,000 bonds. BND provides stability and income, which can help the client achieve their retirement goals.
4. iShares ESG Aware MSCI USA Bond ETF (ESGB): This ETF invests in companies that meet ESG criteria, such as sustainability and social responsibility. It has a low expense ratio of 0.05% and a diversified portfolio of over 1,000 bonds. ESGB provides exposure to companies that are contributing to the transition to a greener economy, which aligns with the client's goals.

Explanation: Equity Funds: VTI and ESGU are both excellent choices for equity funds. VTI provides broad exposure to the U.S. stock market, which is a



Ask me anything...





+ New chat

Query documents ⓘ

Index
canvas-blogs-EN

a Titan Express



Recommendation:

Based on the client's goals and risk profile, I recommend a mix of equity and bond funds. Here are the details: Equity Funds:

1. Vanguard Total Stock Market and a diversified portfolio of finance, which aligns with the client's goals.
2. iShares ESG Aware MSCI USA It has a low expense ratio of 0.04% contributing to the transition.
3. Vanguard Total Bond Market and a diversified portfolio of bonds.
4. iShares ESG Aware MSCI USA It has a low expense ratio of 0.05% contributing to the transition.

Explanation: Equity Funds: VTI and VBR are good starting point for a diversified

Bond Funds: BND and ESGB are both excellent choices for bond funds. BND provides stability and income, which can help the client achieve their retirement goals. ESGB adds an ESG component, which can help the client align their investments with their values.

Risk Considerations: The client has a medium risk tolerance, which means they are willing to take on some risk in order to achieve higher returns.

Create new fine-tuned model ×

Model name

CustomFinanceRecommender

Use only letters, numbers, and underscores up to 32 characters.

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

Ask me anything...



Data Source:

Local upload

Upload files to import

1 file ready to import

Delete all

portfolio-ready-to-train.csv



Import preview

Previewing first 100 rows

[Close preview](#)

[Create dataset](#)

portfolio-ready-to-train.csv

Use first row as header [i](#)

Delete

input	⋮	output	⋮
Based on the information provid...		Fund Allocation: 60% Stable Ho...	
Based on the information provid...		Fund Allocation: 70% FusionDri...	
Based on the information provid...		Fund Allocation: 50% Early Retir...	
Based on the information provid...		Fund Allocation: 40% Child's Ed...	
Based on the information provid...		Fund Allocation: 60% FusionDri...	
Based on the information provid...		Fund Allocation: 50% Early Retir...	

CSV Dataset for Prompt/Completion

[Select](#)[Fine-tune](#)[Analyze](#)

Fine-tune model configurations

Select up to 3 base models

The foundation models you choose will be the base models that are customized to perform your desired language task.

[Learn more about our foundation models.](#)

Select base models

Titan Express

Dolly-7B-Instruct

Falcon-7B-Instruct

Select input column

The input column contains the prompt that is used to train the model.

Input column

input

Select output column

The output column contains the response that you expect the model to generate.

Output column

output

[Fine-tune](#)[Configure model](#)

FinancialData

Only the columns labeled "Input" and "Output" are used to fine-tune the foundation model.

input

A

output

A



Total columns: 2

Total rows: 112

Total cells: 224





Select

Fine-tune

Analyze

II. Model leaderboard



Base model	Training perplexity	Validation perplexity	Training loss	Validation loss	Status	
Dolly-7B-Instruct	14.695	12.510	2.544	2.527	Ready	
Falcon-7B-Instruct	5.984	7.224	2.015	1.977	Ready	
Titan Express	1.000	1.202	0.001	0.184	Ready	





+ New chat

Query documents ⓘ

Index
canvas-blogs-EN

FinancialPortfolio



a

Titan Express



Based on the information provided for the client and their goals and risk profile, recommend funds to invest on, and explain why.

Client Background: A 38-year-old married couple with moderate monthly expenses and a mortgage. They aim to retire comfortably in 20 years and want to support clean energy initiatives. They have a medium risk tolerance.

Goals: To retire comfortably and invest in clean energy.

Risk Profile: Medium

Recommendation:

Fund Allocation: 40% Clean Energy Impact Fund, 30% Early Retirement Portfolio, 30% Global Diversification Portfolio

Customised Results



Based on the information provided, I recommend the following funds to invest in:

Vanguard Total Stock Market Index Fund Admiral Shares: This fund provides broad exposure to the U.S. stock market and has a low expense ratio of 0.04%. It is a good choice for investors with a medium risk tolerance.

iShares ESG Aware MSCI USA ETF: This fund invests in companies that have strong environmental, social, and governance (ESG) practices. It has a low expense ratio of 0.05% and is a good choice for investors who want to support clean energy initiatives.

Vanguard Total International Stock Index Fund Admiral Shares: This fund provides exposure to stocks from around the world and has a low expense ratio of 0.04%. It is a good choice for investors with a medium risk tolerance who want to diversify their portfolio.

Generic Results

Ask me anything...



Thank you!



Please complete the session
survey in the mobile app

Michael Lin

linmicht@amazon.com



© 2024, Amazon Web Services, Inc. or its affiliates. All rights reserved.