



AWS 決定ガイド

AWS 機械学習サービスの選択



AWS 機械学習サービスの選択: AWS 決定ガイド

Copyright © 2025 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Amazon の商標およびトレードドレスは、Amazon のものではない製品またはサービスと関連付けてはならず、また、お客様に混乱を招くような形や Amazon の信用を傷つけたり失わせたりする形で使用することはできません。Amazon が所有しない商標はすべてそれぞれの所有者に所属します。所有者は必ずしも Amazon と提携していたり、関連しているわけではありません。また、Amazon 後援を受けているとはかぎりません。

Table of Contents

決定ガイド 1

序章 1

を理解する 2

考慮する 4

選択 7

使用アイテム 9

Explore 22

リソース 22

ドキュメント履歴 24

..... xxv

AWS 機械学習サービスの選択

作業をサポートする適切な ML サービスとフレームワークを選択する

目的	ニーズに最適な AWS ML サービスを判断するのに役立ちます。
最終更新日	2024 年 5 月 3 日
対象サービス	• Amazon Augmented AI • Amazon CodeGuru • Amazon Comprehend • Amazon DevOps Guru • Amazon Forecast • Amazon Kendra • Amazon Lex • Amazon Personalize • Amazon Polly • Amazon Rekognition • Amazon SageMaker AI • Amazon Textract • Amazon Transcribe • Amazon Translate

序章

最も基本的な機械学習 (ML) は、データから学習し、パターンを特定し、予測を行い、それらの予測に基づいて行動するためのデジタルツールとサービスを提供するように設計されています。現在、ほぼすべての人工知能 (AI) システムは ML を使用して作成されます。ML は、大量のデータを使用して決定ロジックを作成および検証します。この決定ロジックは AI モデルの基礎を形成します。

AWS 機械学習サービスが適用されるシナリオは次のとおりです。

- 特定のユースケース — AWS 機械学習サービスは、一般的なユースケースや業界向けの幅広い構築済みのアルゴリズム、モデル、ソリューションを使用して、AI を活用したユースケースをサポートできます。Amazon Personalize、Amazon Kendra、Amazon Monitron など、事前トレーニング済みの 23 のサービスから選択できます。
- 機械学習のカスタマイズとスケーリング — Amazon SageMaker AI は、あらゆるユースケースの ML モデルの構築、トレーニング、デプロイを支援するように設計されています。Amazon SageMaker AI と Amazon Bedrock AWS を使用して、独自の基盤モデルを構築したり、オープンソースの基盤モデルにアクセスしたりできます。
- 特殊なインフラストラクチャへのアクセス — 機械学習ワークフローをさらに柔軟に制御する必要があり、基盤となるインフラストラクチャとリソースを自分で管理 AWS したい場合は、 が提供する ML フレームワークとインフラストラクチャを使用します。

この決定ガイドは、適切な質問を行い、基準とビジネス上の問題を評価し、ニーズに最適なサービスを決定するのに役立ちます。

を理解する

組織が AI と ML テクノロジーを採用し続ける中で、ML AWS サービスを理解して選択することの重要性は継続的な課題です。

AWS は、組織が ML モデルをより迅速かつ簡単に構築、トレーニング、デプロイできるように設計されたさまざまな ML サービスを提供します。これらのサービスは、顧客解約予測、不正検出、画像や音声の認識など、さまざまなビジネス上の問題を解決するために使用できます。

What is it?



Artificial intelligence (AI)

Any technique that enables computers to mimic human intelligence using logic, if-then statements, and machine learning



Machine learning (ML)

A subset of AI that uses machines to search for patterns in data to build logic models automatically



Classification AI and Predictive AI

A subset of ML that recognizes patterns to identify something (Classification AI) or predicts future trends based on statistical patterns and historical data (Predictive AI)



Generative AI

A subset of DL that can create new content and ideas powered by large, pretrained models called foundation models (FMs)

ML AWS サービスを深く掘り下げる前に、AI と ML の関係を見てみましょう。

- 人工知能とは、これまでヒューマンインテリジェンスを必要としていたタスクをレプリケートできるあらゆるシステムを記述する方法です。ほとんどの AI ユースケースは確率的な結果を探しています。人間の判断と同様に、高い確実性で予測や決定を行います。
- 現在、ほぼすべての AI システムは機械学習を使用して作成されます。ML は大量のデータを使用して、モデルと呼ばれる決定ロジックを作成および検証します。
- 分類 AI は、何かを識別するためのパターンを認識する ML のサブセットです。予測 AI は、履歴データの統計パターンに基づいて将来の傾向を予測する ML のサブセットです。
- 最後に、生成 AI は、会話、ストーリー、画像、動画、音楽などの新しいコンテンツやアイデアを作成できる深層学習のサブセットです。生成 AI は、Foundation Models または FMs と呼ばれる膨大なデータに基づいて事前トレーニングされた非常に大規模なモデルを利用しています。[Amazon Bedrock](#) は、生成 AI アプリケーションを構築およびスケーリングするための高性能 FMs を選択できるフルマネージドサービスです。[Amazon Q Developer](#) および [Amazon Q Business](#) は、特定のユースケース向けの生成 AI を活用したアシスタントです。

このガイドは、主に分類 AI および予測 AI 機械学習カテゴリのサービスを対象とするように設計されています。

さらに、は、高性能の ML トレーニングと推論のために、専用の高速ハードウェア AWS を提供します。

- [Amazon EC2 P5](#) インスタンスには、機械学習のトレーニングタスクと推論タスクの両方に適した NVIDIA H100 Tensor Core GPUs が搭載されています。[Amazon EC2 G5](#) インスタンスには、最大 8 つの NVIDIA A10G Tensor Core GPUs と第 2 世代 AMD EPYC プロセッサが搭載されており、グラフィックスを多用するさまざまなユースケースや機械学習のユースケースに対応します。
- [AWS Trainium](#)は、100B を超えるパラメータモデルの深層学習 (DL) トレーニング専用の第 2 世代 ML アクセラレーター AWS です。
- [AWS Inferentia 2 ベースの Amazon EC2 Inf2 インスタンス](#)は、DL および生成 AI 推論アプリケーション向けに、Amazon EC2 で低コストで高いパフォーマンスを実現するように設計されています。

考慮する

ML AWS サービスでビジネス上の問題を解決する場合、いくつかの主要な基準を考慮すると、成功を確実にするのに役立ちます。次のセクションでは、ML サービスを選択する際に考慮すべき重要な基準をいくつか概説します。

Problem definition

問題定義

ML ライフサイクルの最初のステップは、ビジネス上の問題をフレーム化することです。さまざまなサービスが異なる問題に対処するように設計されているため、解決しようとしている問題を理解することは、適切な AWS ML サービスを選択する上で不可欠です。また、ML がビジネス問題に最適かどうかを判断することも重要です。

ML が最適であると判断したら、まず、さまざまな専用 AWS AI サービス (音声、ビジョン、ドキュメントなど) から選択できます。

Amazon SageMaker AI は、独自のモデルを構築およびトレーニングする必要がある場合にフルマネージドインフラストラクチャを提供します。AWS は、高度にカスタマイズされた特殊な ML モデルを必要とする場合に、高度な ML フレームワークとインフラストラクチャの選択肢を提供します。AWS は、生成 AI を使用して新しいアプリケーションを構築するための一般的な基盤モデルも幅広く提供しています。

ML algorithm

ML アルゴリズム

解決しようとしているビジネス問題の ML アルゴリズムを選択するかどうかは、作業しているデータの種類と期待される結果によって異なります。以下の情報では、主要な AWS AI/ML サービスカテゴリごとにアルゴリズムを使用する方法について説明します。

- 専用 AI サービス: これらのサービスは、特定のタスク用に最適化された事前トレーニング済みのモデルであるため、ML アルゴリズムをカスタマイズする機能は限られています。通常、入力データと一部のパラメータをカスタマイズできますが、基盤となる ML モデルにアクセスしたり、独自のモデルを構築したりすることはできません。
- Amazon SageMaker AI: このサービスは、ML アルゴリズムを最も柔軟に制御します。SageMaker AI を使用して、独自のアルゴリズムとフレームワークを使用してカスタムモデルを構築したり、が提供する構築済みのモデルとアルゴリズムを使用したりできます AWS。これにより、ML プロセスに対する高度なカスタマイズと制御が可能になります。
- 低レベルの ML フレームワークとインフラストラクチャ: これらのサービスは、ML アルゴリズムを最も柔軟に制御します。これらのサービスを使用して、独自のアルゴリズムとフレームワークを使用して高度にカスタマイズされた ML モデルを構築できます。ただし、これらのサービスを使用するには重要な ML の専門知識が必要であり、すべてのユースケースで実現できるとは限りません。

Security

セキュリティ

VPC にプライベートエンドポイントが必要な場合、オプションは使用している ML AWS サービスのレイヤーによって異なります。具体的には次のとおりです。

- 専用 AI サービス: ほとんどの専用 AI サービスは、現在 VPCs のプライベートエンドポイントをサポートしていません。ただし、Amazon Rekognition Custom Labels と Amazon Comprehend Custom には VPC エンドポイントを使用してアクセスできます。
- Core AI サービス: Amazon Translate、Amazon Transcribe、Amazon Comprehend はすべて VPC エンドポイントをサポートしています。
- Amazon SageMaker AI: SageMaker AI は VPC エンドポイントの組み込みサポートを提供するため、トレーニング済みモデルを VPC 内からのみアクセス可能なエンドポイントとしてデプロイできます。
- 下位レベルの ML フレームワークとインフラストラクチャ: Amazon EC2 インスタンスまたは VPC 内のコンテナにモデルをデプロイし、ネットワーク設定を完全に制御できます。

Latency

レイテンシー

Amazon Rekognition や Amazon Transcribe などの高レベルの AI サービスは、さまざまなユースケースに対応し、速度面で高いパフォーマンスを提供するように設計されています。ただし、特定のレイテンシー要件を満たしていない場合があります。

低レベルの ML フレームワークとインフラストラクチャを使用している場合は、Amazon SageMaker AI を活用することをお勧めします。このオプションは、フルマネージドサービスと最適化されたデプロイオプションにより、通常、カスタムモデルを構築するよりも高速です。高度に最適化されたカスタムモデルは SageMaker AI を上回る可能性があります、構築するにはかなりの専門知識とリソースが必要です。

Accuracy

精度

ML AWS サービスの精度は、必要な特定のユースケースとカスタマイズのレベルによって異なります。Amazon Rekognition などの高レベルの AI サービスは、特定のタスク用に最適化され、多くのユースケースで高い精度を提供する事前トレーニング済みのモデルに基づいて構築されています。

場合によっては、カスタム ML モデルを構築およびトレーニングするためのより柔軟でカスタマイズ可能なプラットフォームを提供する Amazon SageMaker AI を使用することを選択できます。独自のモデルを構築することで、事前トレーニング済みのモデルで可能な精度よりもさらに高い精度を実現できる場合があります。

TensorFlow や Apache MXNet などの ML フレームワークとインフラストラクチャを使用して、特定のユースケースで可能な限り高い精度を提供する高度にカスタマイズされたモデルを構築することもできます。

AWS and responsible AI

AWS と責任ある AI

AWS は、開発プロセスの各段階で責任ある AI を念頭に置いて基盤モデル (FMs) を構築します。設計、開発、デプロイ、運用全体で、次のようなさまざまな要因を考慮します。

1. 精度 (概要が基盤となるドキュメントとどの程度一致しているか、伝記が事実上正しいかどうか)
2. 公平性 (出力が属性グループを同様に扱うかどうか)

3. 知的財産と著作権に関する考慮事項
4. 適切な使用法 (法的助言、診断、または違法行為のためにユーザーリクエストをフィルタリングする)
5. 毒性 (ヘイトスピーチ、冒涇、侮辱)
6. プライバシー (個人情報と顧客プロンプトの保護)

AWS は、これらの問題に対処するソリューションを、トレーニングデータの取得に使用されるプロセス、FM 自体、およびユーザープロンプトと出力の事前処理に使用されるテクノロジーに構築します。

選択

ML サービスオプションを評価する基準がわかったので、組織のニーズに適した AWS ML サービスを選択する準備が整いました。次の表は、どの ML サービスがどの状況に最適化されているかを示しています。ユースケースに最適な AWS ML サービスを決定するのに役立ちます。

Categories	いつ使用しますか？	何に最適化されていますか？	関連する AI/ML サービスまたは環境
特定のユースケース これらの人工知能サービスは、特定のニーズを満たすことを目的としています。これには、パーソナライゼーション、予測、異常検出、音声文字起こしなどが含まれます。これらはサービスとして提供されるため、ML の専門知識を必要とせずにアプリケーションに埋め込むことができます。	特定の構築済みの機能をアプリケーションに統合 AWS する必要がある場合は、が提供する AI サービスを使用します。広範なカスタマイズや機械学習の専門知識は必要ありません。これらのサービスは使いやすく、コーディングや設定をあまり必要としないように設計されています。	これらのサービスは使いやすく、コーディング、設定、ML の専門知識をあまり必要としないように設計されています。	Amazon Augmented AI Amazon CodeGuru Amazon Comprehend Amazon Comprehend Medical Amazon DevOps Guru Amazon Forecast Amazon Kendra Amazon Lex

Categories	いつ使用しますか？	何に最適化されていますか？	関連する AI/ML サービスまたは環境
			Amazon Personalize Amazon Polly Amazon Rekognition Amazon Textract Amazon Transcribe Amazon Translate
ML サービス これらのサービスは、コア AI サービスが提供する構築済みの機能を超えるカスタマイズされた機械学習モデルまたはワークフローを開発するために使用できます。	これらのサービスは、コア AI サービスが提供する構築済みの機能を超える、よりカスタマイズされた機械学習モデルやワークフローが必要な場合に使用します。	これらのサービスは、カスタム機械学習モデルの構築とトレーニング、複数のインスタンスまたは GPU クラスターでの大規模なトレーニング、機械学習モデルのデプロイの制御の強化、リアルタイム推論、end-to-end のワークフローの構築に最適化されています。	Amazon SageMaker AI Amazon SageMaker AI JumpStart SageMaker AI Studio SageMaker AI キャンバス SageMaker AI Studio ラボ SageMaker AI Ground Truth での PyTorch AWS Apache MxNet Hugging Face での TensorFlow AWS

Categories	いつ使用しますか？	何に最適化されていますか？	関連する AI/ML サービスまたは環境
インフラストラクチャ 本番環境に機械学習をデプロイするには、Amazon が AWS で構築されたシリコンで有効にする費用対効果の高いインフラストラクチャが必要です。	モデルのトレーニングコストを最小限に抑え、クラウドで推論を実行する必要がある場合に使用します。	機械学習の費用対効果の高いデプロイをサポートするために最適化されています。	AWS Trainium AWS Inferentia および Inferentia2 Amazon SageMaker AI HyperPod
ツールと関連サービス これらのツールおよび関連サービスは、機械学習のデプロイを容易にするように設計されています。	これらのサービスとツールは、クラウドでの深層学習を加速し、Amazon マシンイメージ、Docker イメージ、エンティティ解決を提供するように設計されています。	クラウドでの深層学習を加速できるように最適化されています。	AWS Deep Learning AMIs AWS 深層学習コンテナ AWS エンティティの解決

使用アイテム

AWS ML サービスを選択する際に適用する必要がある基準を明確に理解できたので、ビジネスニーズに最適な AWS AI/ML サービス (複数可) を選択できます。

選択したサービス (複数可) の使用方法と詳細については、各サービスの仕組みを調べるための 3 つのパスセットが用意されています。最初の一連のパスでは、Amazon Comprehend、Amazon Textract、Amazon Translate、Amazon Lex、Amazon Polly、Amazon Rekognition、Amazon Transcribe の使用を開始するための詳細なドキュメント、実践的なチュートリアル、リソースを提供します。

Amazon Comprehend

- Amazon Comprehend の使用を開始する

Amazon Comprehend コンソールを使用して、非同期エンティティ検出ジョブを作成して実行します。

[チュートリアルを開始方法」](#)

- Amazon Comprehend を使用してテキストでインサイトを分析する

Amazon Comprehend を使用してテキストを分析し、テキストからインサイトを引き出す方法について説明します。

[チュートリアルを開始方法」](#)

- Amazon Comprehend の料金

Amazon Comprehend の料金と例について説明します。

[ガイドを詳しく見る »](#)

Amazon Textract

- Amazon Textract の開始方法

Amazon Textract をフォーマットされたテキストでを使用して、互いに近い単語や単語行を検出する方法、および関連するテキスト、テーブル、キーと値のペア、選択要素などの項目についてドキュメントを分析する方法について説明します。

[ガイドを詳しく見る »](#)

- Extract text and structured data with Amazon Textract

Amazon Textract を使用してドキュメントからテキストと構造化データを抽出する方法について説明します。

[チュートリアルを開始方法」](#)

このエピソードで Amazon Textract の詳細を確認し、 で時間を費やして AWS Management Console、サービス APIs を最大限に活用する方法を理解するのに役立つコードサンプルを確認します。

[動画を見る »](#)

Amazon Translate

- コンソールを使用した Amazon Translate の開始方法

Amazon Translate の使用を開始する最も簡単な方法は、コンソールを使って何らかのテキストを翻訳することです。コンソールを使用して最大 10,000 文字を翻訳する方法について説明します。

[ガイドを詳しく見る »](#)

- クラウド内の言語間でテキストを翻訳する

このチュートリアル例では、国際的な卸売製造会社の一部として、顧客が製品について何を言っているのかを、現地市場言語 - フランス語でのレビューで理解する必要があります。

[チュートリアルの開始方法」](#)

- Amazon Translate の料金

12 か月間、1 か月あたり 200 万文字を提供する 無料利用枠を含む Amazon Translate の料金をご覧ください。

[ガイドを詳しく見る »](#)

Amazon Lex

- Amazon Lex V2 デベロッパーガイド

Amazon Lex V2 の開始方法、仕組み、料金情報について説明します。

[ガイドを詳しく見る »](#)

- Amazon Lex の概要 Amazon Lex 会話サービスを紹介し、ボットを作成してさまざまなチャットサービスにデプロイする方法を示す例を紹介します。

[コースを受講する »](#) (サインインが必要)

- 会話体験における生成 AI の探索

会話エクスペリエンスでの生成 AI の使用について説明します。

[ブログを読む »](#)

Amazon Polly

- Amazon Polly とは

テキストをリアルな音声に変換し、カスタマーエンゲージメントとアクセシビリティを向上させるアプリケーションの開発に使用できるクラウドサービスの完全な概要について説明します。

[ガイドを詳しく見る »](#)

- Amazon Polly を使用して発声されるテキストを強調表示する

書籍、ウェブサイト、ブログ、その他のデジタルエクスペリエンスのオーディオにビジュアル機能を追加するために話されているテキストを強調するアプローチを紹介します。

[ブログを読む »](#)

- Amazon Polly で同じ TTS 音声ペルソナを使用して、複数の言語でコンテンツの音声を作成する

Neural Text-to-Speech (NTTS) について説明し、サポートされている言語でさまざまな異なるスピーカーを提供する、利用可能な音声の幅広いポートフォリオがどのように機能するかについて説明します。

[ブログを読む »](#)

Amazon Rekognition

- Amazon Rekognition とは

このサービスを使用して、アプリケーションにイメージ分析とビデオ分析を追加する方法について説明します。

[ガイドを詳しく見る »](#)

- ハンズオン Rekognition: 自動画像とビデオ分析

顔認識がストリーミングビデオとどのように連携するか、コード例とキーポイントをセルフガイドのペースで説明します。

[チュートリアルを開始方法」](#)

- Amazon Rekognition FAQs

Amazon Rekognition の基本と、深層学習を改善し、アプリケーションを視覚的に分析するのに役立つ方法について説明します。

[FAQsを読む](#)

Amazon Transcribe

- Amazon Transcribe とは

ML を使用して音声をテキストに変換する AWS 自動音声認識サービスについて説明します。このサービスをスタンドアロン文字起こしとして使用したり、任意のアプリケーションに speech-to-text 機能を追加する方法について説明します。

[ガイドを詳しく見る »](#)

- Amazon Transcribe の料金

カスタム言語モデルオプションや Amazon Transcribe 無料利用枠など、AWS pay-as-you-go の文字起こしについて説明します。

[ガイドを詳しく見る »](#)

- Amazon Transcribe でオーディオトランスクリプトを作成する

Amazon Transcribe を使用して、ニーズに照らしてテストするための実際のユースケースシナリオを使用して、録音されたオーディオファイルのテキストトランスクリプトを作成する方法について説明します。

[チュートリアルを開始方法」](#)

- Amazon Transcribe ストリーミングアプリを構築する

ライブオーディオをリアルタイムで記録、文字起こし、翻訳し、結果を直接 E メールで送信するアプリを構築する方法について説明します。

[ガイドを詳しく見る »](#)

2 番目の AI/ML AWS のサービス パッケージでは、Amazon SageMaker AI ファミリーのサービスを開始するための詳細なドキュメント、実践的なチュートリアル、およびリソースを提供します。

SageMaker AI

- Amazon SageMaker AI の仕組み

機械学習の概要と SageMaker AI の仕組みについて説明します。

[ガイドを詳しく見る »](#)

- Amazon SageMaker AI の開始方法

Amazon SageMaker AI ドメインに参加して、Amazon SageMaker AI Studio と RStudio on SageMaker AI にアクセスする方法について説明します。

[ガイドを詳しく見る »](#)

- Amazon SageMaker AI で Apache Spark を使用する

Apache Spark を使用してデータを前処理し、SageMaker AI を使用してモデルトレーニングとホスティングを行う方法について説明します。

[ガイドを詳しく見る »](#)

- Docker コンテナを使用してモデルを構築する

Amazon SageMaker AI がビルドタスクとランタイムタスクに Docker コンテナを幅広く活用する方法について説明します。組み込みアルゴリズム用の構築済み Docker イメージと、トレーニングと推論に使用されるサポートされている深層学習フレームワークをデプロイする方法について説明します。

[ガイドを詳しく見る »](#)

- 機械学習フレームワークと言語

Amazon SageMaker AI Python SDK を使用して SageMaker AI の使用を開始する方法について説明します。Amazon SageMaker

[ガイドを詳しく見る »](#)

SageMaker AI Autopilot

- 表形式データの Amazon SageMaker AI Autopilot 実験を作成する

Amazon SageMaker AI Autopilot 実験を作成して、表形式のデータセットでさまざまなモデル候補を探索、前処理、トレーニングする方法について説明します。

[ガイドを詳しく見る »](#)

- 機械学習モデルを自動的に作成する

Amazon SageMaker AI Autopilot を使用して ML モデルを自動的に構築、トレーニング、チューニングし、モデルをデプロイして予測を行う方法について説明します。

[チュートリアルの開始方法」](#)

- これらのサンプルノートブックを使用して Amazon SageMaker AI Autopilot によるモデリングを試す

ダイレクトマーケティング、顧客解約予測、独自のデータ処理コードを Amazon SageMaker AI Autopilot に持ち込む方法に関するサンプルノートブックについて説明します。

[ガイドを詳しく見る »](#)

SageMaker AI Canvas

- Amazon SageMaker AI Canvas の使用を開始する

SageMaker AI Canvas の使用を開始する方法について説明します。

[ガイドを詳しく見る »](#)

- コードを記述せずに機械学習予測を生成する

このチュートリアルでは、Amazon SageMaker AI Canvas を使用して ML モデルを構築し、1 行のコードを記述せずに正確な予測を生成する方法について説明します。

[チュートリアルの開始方法」](#)

- SageMaker AI Canvas の詳細

SageMaker AI Canvas とその視覚的なノーコード ML 機能について詳しく説明します。

[ブログを読む »](#)

- Amazon SageMaker AI Canvas を使用して最初の ML モデルを作成する

Amazon SageMaker AI Canvas を使用して、新しい製品やサービスの E メールキャンペーンに基づいて顧客維持を評価する ML モデルを作成する方法について説明します。

[ラボの使用を開始する」](#)

SageMaker AI Data Wrangler

- Amazon SageMaker AI Data Wrangler の開始方法

SageMaker AI Data Wrangler をセットアップし、既存のサンプルデータセットを使用したチュートリアルを提供する方法について説明します。

[ガイドを詳しく見る »](#)

- 最小限のコードで機械学習用のトレーニングデータを準備する

Amazon SageMaker AI Data Wrangler を使用して ML のデータを準備する方法について説明します。

[チュートリアルを開始方法」](#)

- SageMaker AI Data Wrangler デイープダイブワークショップ

データセットに適切な分析タイプを適用して異常や問題を検出し、導出された結果/インサイトを使用してデータセットの変換の過程で修復アクションを策定し、SageMaker AI Data Wrangler が提供するクイックモデリングオプションを使用して変換の適切な選択と順序をテストする方法について説明します。

[ワークショップの開始方法 »](#)

SageMaker AI Ground Truth

- Amazon Ground Truth の開始方法

コンソールを使用してラベル付けジョブを作成し、パブリックまたはプライベートワークフォースを割り当て、ラベル付けジョブをワークフォースに送信する方法について説明します。ラベル付けジョブの進行状況をモニタリングする方法について説明します。

[ガイドを詳しく見る »](#)

- Machine Learningのトレーニングデータのラベル付け

Amazon SageMaker AI Ground Truth でラベリングジョブをセットアップして、ML モデルのトレーニングデータに注釈を付ける方法について説明します。

[チュートリアルを開始方法」](#)

- Amazon Ground Truth Plus の開始方法 Amazon SageMaker AI Ground Truth Plus プロジェクトを開始し、ラベルを確認し、SageMaker AI Ground Truth Plus の前提条件を満たすために必要なステップを完了する方法について説明します。

[ガイドを詳しく見る »](#)

- Amazon Ground Truth Watch の使用を開始します。SageMaker AI Ground Truth コンソールを使用して、数分でデータのラベル付けを開始する方法を説明します。

[動画を見る »](#)

- Amazon SageMaker AI Ground Truth Plus – コードや社内リソースなしでトレーニングデータセットを作成する

エキスパートワークフォースを使用して高品質のトレーニングデータセットを迅速に提供し、コストを最大 40% 削減するターンキーサービスである Ground Truth Plus について説明します。

[ブログを読む »](#)

SageMaker AI JumpStart

- SageMaker AI JumpStart で機械学習を開始する

一般的なユースケース用にインフラストラクチャを設定する SageMaker AI JumpStart ソリューションテンプレートと、SageMaker AI を使用した機械学習用の実行可能サンプルノートブックについて説明します。

[ガイドを詳しく見る »](#)

- Amazon SageMaker AI JumpStart を使用して機械学習プロジェクトをすばやく開始する

Amazon SageMaker AI JumpStart が提供する事前トレーニング済みモデルと事前構築済みソリューションを使用して ML プロジェクトを高速追跡する方法について説明します。その後、Amazon SageMaker AI Studio ノートブックを使用して選択したモデルをデプロイできます。

[チュートリアルの開始方法」](#)

- この Immersion Day ワークショップで Amazon SageMaker AI JumpStart を実践する

Amazon SageMaker AI Data Wrangler、Autopilot、Jumpstart で見つかったローコード ML 機能を使用して、より迅速に実験し、非常に正確なモデルを本番環境に導入する方法について説明します。

[ワークショップの開始方法 »](#)

SageMaker AI Pipelines

- Amazon SageMaker AI Pipelines の開始方法

SageMaker AI ジョブを管理およびデプロイするend-to-endのワークフローを作成する方法について説明します。SageMaker AI Pipelines には SageMaker AI Python SDK 統合が付属しているため、Python ベースのインターフェイスを使用してパイプラインの各ステップを構築できます。

[ガイドを詳しく見る »](#)

- 機械学習ワークフローを自動化する

Amazon SageMaker AI Pipelines、Amazon SageMaker AI Model Registry、Amazon SageMaker Amazon SageMaker AI Clarify を使用してend-to-endの機械学習 (ML) ワークフローを作成および自動化する方法について説明します。

[チュートリアルの開始方法」](#)

- Amazon SageMaker AI Pipelines で完全に自動化された ML ワークフローを作成する方法

Amazon SageMaker AI Pipelines について説明します。これは、すべての開発者とデータサイエンティストがアクセスできるように設計された、世界初の ML CI/CD サービスです。SageMaker AI Pipelines は CI/CD パイプラインを ML に導入し、必要なコーディング時間を短縮します。

[動画を見る](#) ◆◆

SageMaker AI Studio

- 機械学習モデルをローカルで構築してトレーニングする

Amazon SageMaker AI Studio ノートブック内で ML モデルをローカルで構築およびトレーニングする方法について説明します。

[チュートリアルを開始方法」](#)

- SageMaker AI Studio と EMR の統合ワークショップ

分散処理を大規模に活用してデータを準備し、その後機械学習モデルをトレーニングする方法について説明します。

[ワークショップを開始方法 »](#)

3 番目の AI/ML AWS のサービス パッケージでは、AWS Trainium、Amazon Titan の使用を開始するための詳細なドキュメント、実践的なチュートリアル AWS Inferentia、リソースを提供します。

AWS Trainium

- AWS Trainium と Amazon EKS を使用した分散トレーニングのスケーリング

クラウドで深層学習モデルをトレーニングするための高性能で費用対効果が高く、非常にスケラブルなプラットフォームを提供するために最適化された専用の ML アクセラレーター AWS Trainiumである を搭載した Amazon EC2 Trn1 インスタンスの一般提供を活用する方法について説明します。

[ブログを読む »](#)

- の概要 AWS Trainium

AWS Trainium 100B 以上のパラメータモデルの深層学習トレーニング用に AWS 構築された第 2 世代機械学習 (ML) アクセラレーターについて説明します。各 Amazon Elastic Compute Cloud (EC2) Trn1 インスタンスは、最大 16 個の AWS Trainium アクセラレーターをデプロイして、クラウドでの深層学習 (DL) トレーニングのための高性能で低コストのソリューションを提供します。

[ガイドを詳しく見る »](#)

- Recommended Trainium Instances

深層学習モデル推論ワークロードに高いパフォーマンスとコスト効率を提供するように AWS Trainium インスタンスを設計する方法について説明します。

[ガイドを詳しく見る »](#)

AWS Inferentia

- の概要 AWS Inferentia

深層学習 (DL) 推論アプリケーションの低コストで高いパフォーマンスを実現する AWS ように、
によってアクセラレーターがどのように設計されているかを理解します。

[ガイドを詳しく見る »](#)

- AWS Inferentiaスループットが 4 倍、レイテンシーが AWS Inferentia 10 倍に向上する 2 つの
ビルド on1

に最適化された AWS Inferentia 2 とは何かを理解し、LLMsと生成 AI 推論のコストを削減しながら、
パフォーマンスを向上させるためにゼロからどのように設計されたかを調べます。

[ブログを読む »](#)

- を使用した機械学習推論 AWS Inferentia

Amazon EC2 Inf1 インスタンスを実行しているノードを使用して Amazon Inf1クラスターを
作成し、(オプションで) サンプルアプリケーションをデプロイする方法について説明しま
す。Amazon EC2 Inf1 インスタンスは AWS Inferentia 、クラウド内で高パフォーマンスと低コ
ストの推論 AWS を提供するために、
によってカスタム構築されたチップを搭載しています。

[ガイドを詳しく見る »](#)

Amazon Titan

- Amazon Titan の概要

Amazon Titan FMs を大規模なデータセットで事前トレーニングし、強力な汎用的なモデルに
する方法について説明します。それらをそのまま、またはプライベートに使用し、大量のデー
タに注釈を付けずに、特定のタスクの独自のデータでカスタマイズする方法について説明しま
す。

[ガイドを詳しく見る »](#)

Explore

- **アーキテクチャ図**

これらのリファレンスアーキテクチャ図は、使用中の AWS AI および ML サービスの例を示しています。

[アーキテクチャ図を調べる »](#)

- **ホワイトペーパー**

AI/ML サービスの選択と使用を開始してベストプラクティスを学ぶのに役立つホワイトペーパーをご覧ください。

[ホワイトペーパーを見る »](#)

- **AWS ソリューション**

AI および ML サービスの一般的なユースケースについて、厳選されたソリューションとアーキテクチャガイダンスをご覧ください。

[ソリューションの詳細 »](#)

リソース

基盤モデル

サポートされている基盤モデルは次のとおりです。

- [Anthropic Claude](#)
- [Cohere コマンドと埋め込み](#)
- [AI21 ラボ Jurassic](#)
- [メタラマ](#)
- [霧 AI](#)
- [Stable Diffusion XL](#)
- [Amazon Titan](#)

Amazon Bedrock を使用すると、さまざまな基盤モデルを試し、データでプライベートにカスタマイズできます。

ユースケースまたは業界固有のサービス

- [Amazon Comprehend Medical](#)
- [Amazon Fraud Detector](#)
- [AWS HealthLake](#)
- [Amazon Lookout for Equipment](#)
- [Amazon Lookout for Metrics](#)
- [Amazon Lookout for Vision](#)
- [Amazon Monitron](#)
- [AWS HealthOmics](#)
- [AWS Panorama](#)

関連するブログ投稿

- [重要な新機能により、Amazon Bedrock を使用して生成 AI アプリケーションの構築とスケーリングを容易にし、優れた結果を達成できます。](#)
- [AWS Amazon SageMaker AI JumpStart に Llama 3 モデルをデプロイするための推論と最低コスト AWS Trainium の提供](#)
- [Amazon SageMaker AI でビジネスに合わせた報酬モデルで顧客満足度を変革する](#)
- [Amazon Personalize が、レイテンシーの低い大規模なアイテムカタログをサポートする新しいレシピをリリース](#)

ドキュメント履歴

次の表に、この決定ガイドの重要な変更点を示します。このガイドの更新に関する通知については、RSS フィードをサブスクライブできます。

変更	説明	日付
マイナーな更新	Amazon Q と Amazon の最新の AI および ML スタックのコンテンツを更新しました。	2024 年 5 月 3 日
初回リリース	決定ガイドの初回リリース。	2023 年 7 月 24 日

翻訳は機械翻訳により提供されています。提供された翻訳内容と英語版の間で齟齬、不一致または矛盾がある場合、英語版が優先します。