



AWS 결정 가이드

AWS 기계 학습 서비스 선택



AWS 기계 학습 서비스 선택: AWS 결정 가이드

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

Amazon의 상표 및 브랜드 디자인은 Amazon 외 제품 또는 서비스와 함께, Amazon 브랜드 이미지를 떨어뜨리거나 고객에게 혼동을 일으킬 수 있는 방식으로 사용할 수 없습니다. Amazon이 소유하지 않은 기타 모든 상표는 Amazon과 제휴 관계이거나 관련이 있거나 후원 관계 여부에 관계없이 해당 소유자의 자산입니다.

Table of Contents

결정 가이드	1
소개	1
이해	2
고려 사항	4
선택	7
사용	9
탐색	21
리소스	21
문서 기록	23
.....	xxiv

AWS 기계 학습 서비스 선택

작업을 지원하는 적절한 ML 서비스 및 프레임워크 선택

용도	필요에 가장 적합한 AWS ML 서비스를 결정하는 데 도움이 됩니다.
최종 업데이트 날짜	2024년 5월 3일
적용 대상 서비스	• 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

소개

가장 기본적인 기계 학습(ML)은 데이터에서 학습하고, 패턴을 식별하고, 예측하고, 이러한 예측에 따라 조치를 취할 수 있는 디지털 도구와 서비스를 제공하도록 설계되었습니다. 오늘날 거의 모든 인공지능(AI) 시스템은 ML을 사용하여 생성됩니다. ML은 대량의 데이터를 사용하여 결정 로직을 생성하고 검증합니다. 이 결정 로직은 AI 모델의 기반을 형성합니다.

AWS 기계 학습 서비스를 적용할 수 있는 시나리오는 다음과 같습니다.

- 특정 사용 사례 - AWS 기계 학습 서비스는 일반적인 사용 사례 및 산업을 위한 광범위한 사전 구축 알고리즘, 모델 및 솔루션을 통해 AI 기반 사용 사례를 지원할 수 있습니다. Personalize, Amazon Kendra, Amazon Monitron 등 23개의 사전 훈련된 서비스를 선택할 수 있습니다.
- 기계 학습 사용자 지정 및 조정 - Amazon SageMaker AI는 모든 사용 사례에 맞게 ML 모델을 구축, 훈련 및 배포할 수 있도록 설계되었습니다. Amazon SageMaker AI 및 Amazon Bedrock을 AWS 통해서 자체 오픈 소스 기본 모델을 빌드하거나 액세스할 수 있습니다.
- 특수 인프라 액세스 - 기계 학습 워크플로를 훨씬 더 유연하게 제어하고 기본 인프라와 리소스를 직접 관리하려는 AWS 경우에서 제공하는 ML 프레임워크 및 인프라를 사용합니다.

이 결정 가이드는 올바른 질문을 하고, 기준 및 비즈니스 문제를 평가하고, 필요에 가장 적합한 서비스를 결정하는 데 도움이 됩니다.

이해

조직이 AI 및 ML 기술을 계속 채택함에 따라 AWS ML 서비스를 이해하고 선택하는 것은 지속적인 과제입니다.

AWS 는 조직이 ML 모델을 더 빠르고 쉽게 빌드, 훈련 및 배포할 수 있도록 설계된 다양한 ML 서비스를 제공합니다. 이러한 서비스를 사용하여 고객 이탈 예측, 사기 탐지, 이미지 및 음성 인식과 같은 다양한 비즈니스 문제를 해결할 수 있습니다.

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)

AWS ML 서비스에 대해 자세히 알아보기 전에 AI와 ML 간의 관계를 살펴보겠습니다.

- 높은 수준에서 인공 지능은 이전에 인간 지능이 필요했던 작업을 복제할 수 있는 모든 시스템을 설명하는 방법입니다. 대부분의 AI 사용 사례는 인간의 판단과 마찬가지로 높은 수준의 확실성을 갖춘 예측 또는 결정을 내리는 확률적 결과를 찾고 있습니다.
- 오늘날 거의 모든 AI 시스템은 기계 학습을 사용하여 생성됩니다. ML은 대량의 데이터를 사용하여 모델이라고 하는 결정 로직을 생성하고 검증합니다.
- 분류 AI는 패턴을 인식하여 무언가를 식별하는 ML의 하위 집합입니다. 예측 AI는 통계 패턴 및 과거 데이터를 기반으로 미래 추세를 예측하는 ML의 하위 집합입니다.
- 마지막으로 생성형 AI는 대화, 스토리, 이미지, 비디오 및 음악과 같은 새로운 콘텐츠와 아이디어를 생성할 수 있는 딥 러닝의 하위 집합입니다. 생성형 AI는 파운데이션 모델 또는 FMs이라는 방대한 데이터 코포라를 기반으로 사전 훈련된 매우 큰 모델로 구동됩니다. [Amazon Bedrock](#)은 생성형 AI 애플리케이션을 구축하고 확장하기 위한 고성능 FMs을 선택할 수 있는 완전관리형 서비스입니다. [Amazon Q Developer](#) 및 [Amazon Q Business](#)are 특정 사용 사례를 위한 생성형 AI 기반 어시스턴트입니다.

이 가이드는 주로 분류 AI 및 예측 AI 기계 학습 범주의 서비스를 다루도록 설계되었습니다.

또한 고성능 ML 훈련 및 추론을 위한 특수 가속 하드웨어를 AWS 제공합니다.

- [Amazon EC2 P5instances](#)에는 기계 학습의 훈련 및 추론 작업에 모두 적합한 NVIDIA H100 Tensor Core GPUs가 탑재되어 있습니다. [Amazon EC2 G5instances](#)는 다양한 그래픽 집약적 및 기계 학습 사용 사례를 위해 최대 8개의 NVIDIA A10G Tensor Core GPUs와 2세대 AMD EPYC 프로세서를 제공합니다.
- [AWS Trainium](#)은 100B AWS 개 이상의 파라미터 모델의 딥 러닝(DL) 훈련을 위해 특별히 구축된 2세대 ML 액셀러레이터입니다.
- [AWS Inferentia 2 기반 Amazon EC2 Inf2 인스턴스](#)는 DL 및 생성형 AI 추론 애플리케이션을 위해 Amazon EC2에서 최저 비용으로 고성능을 제공하도록 설계되었습니다.

고려 사항

AWS ML 서비스의 비즈니스 문제를 해결할 때 몇 가지 주요 기준을 고려하면 성공을 보장하는 데 도움이 될 수 있습니다. 다음 섹션에서는 ML 서비스를 선택할 때 고려해야 할 몇 가지 주요 기준을 간략하게 설명합니다.

Problem definition

문제 정의

ML 수명 주기의 첫 번째 단계는 비즈니스 문제를 구성하는 것입니다. 다양한 서비스가 다양한 문제를 해결하도록 설계되었으므로 해결하려는 문제를 이해하는 것은 올바른 AWS ML 서비스를 선택하는 데 필수적입니다. ML이 비즈니스 문제에 가장 적합한지 확인하는 것도 중요합니다.

ML이 가장 적합하다고 판단되면 다양한 목적별 AWS AI 서비스(예: 음성, 비전 및 문서) 중에서 선택하여 시작할 수 있습니다.

Amazon SageMaker AI는 자체 모델을 빌드하고 훈련해야 하는 경우 완전 관리형 인프라를 제공합니다.는 고도로 사용자 지정되고 특수화된 ML 모델이 필요한 경우에 적합한 다양한 고급 ML 프레임워크 및 인프라 선택을 AWS 제공합니다. AWS 또한 생성형 AI를 사용하여 새 애플리케이션을 빌드하기 위한 널리 사용되는 다양한 파운데이션 모델을 제공합니다.

ML algorithm

ML 알고리즘

해결하려는 비즈니스 문제에 대한 ML 알고리즘을 선택하는 것은 작업 중인 데이터의 유형과 원하는 결과에 따라 달라집니다. 다음 정보는 각 주요 AWS AI/ML 서비스 범주가 알고리즘으로 작업할 수 있도록 지원하는 방법을 간략하게 설명합니다.

- **전문화된 AI 서비스:** 이러한 서비스는 특정 작업에 최적화된 사전 훈련된 모델이므로 ML 알고리즘을 사용자 지정하는 제한된 기능을 제공합니다. 일반적으로 입력 데이터와 일부 파라미터를 사용자 지정할 수 있지만 기본 ML 모델 또는 자체 모델 빌드 기능에는 액세스할 수 없습니다.
- **Amazon SageMaker AI:** 이 서비스는 ML 알고리즘을 가장 유연하게 제어합니다. SageMaker AI를 사용하여 자체 알고리즘 및 프레임워크를 사용하여 사용자 지정 모델을 빌드하거나에서 제공하는 사전 빌드된 모델 및 알고리즘을 사용할 수 있습니다 AWS. 이를 통해 ML 프로세스에 대한 높은 수준의 사용자 지정 및 제어가 가능합니다.
- **하위 수준 ML 프레임워크 및 인프라:** 이러한 서비스는 ML 알고리즘을 가장 유연하게 제어합니다. 이러한 서비스를 사용하여 자체 알고리즘 및 프레임워크를 사용하여 고도로 사용자 지정된 ML 모델을 구축할 수 있습니다. 그러나 이러한 서비스를 사용하려면 상당한 ML 전문 지식이 필요하며 모든 사용 사례에 적합하지 않을 수 있습니다.

Security

[Security]

VPC에 프라이빗 엔드포인트가 필요한 경우 사용 중인 AWS ML 서비스 계층에 따라 옵션이 달라집니다. 다음이 포함됩니다.

- **전문화된 AI 서비스:** 대부분의 전문화된 AI 서비스는 현재 VPCs의 프라이빗 엔드포인트를 지원하지 않습니다. 그러나 VPC 엔드포인트를 사용하여 Amazon Rekognition Custom Labels 및 Amazon Comprehend Custom에 액세스할 수 있습니다.
- **코어 AI 서비스:** Amazon Translate, Amazon Transcribe 및 Amazon Comprehend는 모두 VPC 엔드포인트를 지원합니다.
- **Amazon SageMaker AI:** SageMaker AI는 VPC 엔드포인트에 대한 기본 지원을 제공하므로 훈련된 모델을 VPC 내에서만 액세스할 수 있는 엔드포인트로 배포할 수 있습니다.
- **하위 수준 ML 프레임워크 및 인프라:** Amazon EC2 인스턴스 또는 VPC 내 컨테이너에 모델을 배포하여 네트워킹 구성을 완벽하게 제어할 수 있습니다.

Latency

Latency

Amazon Rekognition 및 Amazon Transcribe와 같은 상위 수준 AI 서비스는 다양한 사용 사례를 처리하고 속도 측면에서 높은 성능을 제공하도록 설계되었습니다. 그러나 특정 지연 시간 요구 사항을 충족하지 못할 수 있습니다.

하위 수준 ML 프레임워크 및 인프라를 사용하는 경우 Amazon SageMaker AI를 활용하는 것이 좋습니다. 이 옵션은 완전 관리형 서비스와 최적화된 배포 옵션으로 인해 일반적으로 사용자 지정 모델을 빌드하는 것보다 빠릅니다. 고도로 최적화된 사용자 지정 모델은 SageMaker AI보다 성능이 뛰어나지만 구축하려면 상당한 전문 지식과 리소스가 필요합니다.

Accuracy

정확도

AWS ML 서비스의 정확도는 특정 사용 사례와 필요한 사용자 지정 수준에 따라 달라집니다. Amazon Rekognition과 같은 상위 수준 AI 서비스는 특정 작업에 최적화되고 많은 사용 사례에서 높은 정확도를 제공하는 사전 훈련된 모델을 기반으로 합니다.

경우에 따라 사용자 지정 ML 모델을 구축하고 훈련하기 위한 보다 유연하고 사용자 지정 가능한 플랫폼을 제공하는 Amazon SageMaker AI를 사용하도록 선택할 수 있습니다. 자체 모델을 구축하면 사전 훈련된 모델보다 훨씬 더 높은 정확도를 달성할 수 있습니다.

TensorFlow 및 Apache MXNet과 같은 ML 프레임워크 및 인프라를 사용하여 특정 사용 사례에 가장 높은 정확도를 제공하는 고도로 사용자 지정된 모델을 구축하도록 선택할 수도 있습니다.

AWS and responsible AI

AWS 및 책임 있는 AI

AWS는 개발 프로세스의 각 단계에서 책임감 있는 AI를 염두에 두고 파운데이션 모델(FMs)을 구축합니다. 설계, 개발, 배포 및 운영 전반에 걸쳐 다음과 같은 다양한 요소를 고려합니다.

1. 정확도(요약이 기본 문서와 얼마나 일치하는지, 바이오그래피가 사실적으로 올바른지 여부)
2. 공정성(출력이 인구통계 그룹을 비슷하게 처리하는지 여부)
3. 지적 재산 및 저작권 고려 사항
4. 적절한 사용(법적 조언, 의료 진단 또는 불법 활동에 대한 사용자 요청 필터링)
5. 유해성(증오 발언, 비속어 및 모욕)
6. 개인 정보 보호(개인 정보 및 고객 프롬프트 보호)

AWS는 이러한 문제를 해결하는 솔루션을 훈련 데이터를 획득하는 데 사용되는 프로세스, FMs 자체, 사용자 프롬프트 및 프로세스 후 출력을 사전 처리하는 데 사용되는 기술에 구축합니다.

선택

이제 ML 서비스 옵션을 평가하는 기준을 알았으므로 조직의 요구 사항에 적합한 AWS ML 서비스를 선택할 준비가 되었습니다. 다음 표에서는 어떤 ML 서비스가 어떤 상황에 최적화되어 있는지 보여줍니다. 이를 사용하여 사용 사례에 가장 적합한 AWS ML 서비스를 확인할 수 있습니다.

Categories	언제 사용하나요?	무엇에 최적화되어 있나요?	관련 AI/ML 서비스 또는 환경
특정 사용 사례 이러한 인공 지능 서비스는 특정 요구 사항을 충족하기 위한 것입니다. 여기에는 개인화, 예측, 이상 탐지, 음성 트랜스크립션 등이 포함됩니다. 서비스로 제공되므로 ML 전문 지식 없이도 애플리케이션에 포함할 수 있습니다.	광범위한 사용자 지정 또는 기계 학습 전문 지식 없이 특정 사전 구축 기능을 애플리케이션에 통합해야 하는 AWS 경우에서 제공하는 AI 서비스를 사용합니다. 이러한 서비스는 사용하기 쉽도록 설계되었으며 코딩이나 구성이 많이 필요하지 않습니다.	이러한 서비스는 사용하기 쉽도록 설계되었으며 코딩, 구성 또는 ML 전문 지식이 많이 필요하지 않습니다.	Amazon Augmented AI Amazon CodeGuru Amazon Comprehend Amazon Comprehend Medical Amazon DevOps Guru Amazon Forecast Amazon Kendra Amazon Lex Amazon Personalize Amazon Polly Amazon Rekognition Amazon Textract Amazon Transcribe Amazon Translate

Categories	언제 사용하나요?	무엇에 최적화되어 있나요?	관련 AI/ML 서비스 또는 환경
ML 서비스 이러한 서비스는 핵심 AI 서비스에서 제공하는 사전 구축된 기능을 넘어서는 사용자 지정 기계 학습 모델 또는 워크플로를 개발하는 데 사용할 수 있습니다.	코어 AI 서비스에서 제공하는 사전 구축된 기능을 넘어서는 더 많은 사용자 지정 기계 학습 모델 또는 워크플로가 필요한 경우 이러한 서비스를 사용합니다.	이러한 서비스는 사용자 지정 기계 학습 모델 구축 및 훈련, 여러 인스턴스 또는 GPU 클러스터에 대한 대규모 훈련, 기계 학습 모델 배포에 대한 더 많은 제어, 실시간 추론, end-to-end 워크플로 구축에 최적화되어 있습니다.	Amazon SageMaker AI Amazon SageMaker AI JumpStart SageMaker AI Studio SageMaker AI 캔버스 SageMaker AI Studio Lab SageMaker AI Ground Truth 의 PyTorch AWS Apache MxNet 안개 얼굴 의 TensorFlow AWS
인프라 프로덕션 환경에서 기계 학습을 배포하려면 Amazon이 구축 AWS 된 실리콘으로 지원하는 비용 효율적인 인프라가 필요합니다.	모델 훈련 비용을 최소화하고 클라우드에서 추론을 실행해야 할 때 사용합니다.	기계 학습의 비용 효율적인 배포를 지원하도록 최적화되었습니다.	AWS Trainium AWS Inferentia 및 Inferentia2 Amazon SageMaker AI HyperPod

Categories	언제 사용하나요?	무엇에 최적화되어 있나요?	관련 AI/ML 서비스 또는 환경
도구 및 관련 서비스 이러한 도구 및 관련 서비스는 기계 학습을 쉽게 배포할 수 있도록 설계되었습니다.	이러한 서비스 및 도구는 클라우드에서 딥 러닝을 가속화하여 Amazon 머신 이미지, 도커 이미지 및 개체 해상도를 제공하도록 설계되었습니다.	클라우드에서 딥 러닝을 가속화하는 데 최적화되었습니다.	AWS Deep Learning AMIs AWS 딥 러닝 컨테이너 AWS 개체 해결

사용

이제 AWS ML 서비스를 선택할 때 적용해야 하는 기준을 명확하게 이해했으므로 비즈니스 요구 사항에 최적화된 AWS AI/ML 서비스(들)를 선택할 수 있습니다.

선택한 서비스(들)를 사용하고 자세히 알아보는 방법을 알아보기 위해 각 서비스의 작동 방식을 탐색하는 세 가지 경로 세트를 제공했습니다. 첫 번째 경로 세트는 Amazon Comprehend, Amazon Textract, Amazon Translate, Amazon Lex, Amazon Polly, Amazon Rekognition 및 Amazon Transcribe를 시작하기 위한 심층 설명서, 실습 자습서 및 리소스를 제공합니다.

Amazon Comprehend

- Amazon Comprehend 시작하기

Amazon Comprehend 콘솔을 사용하여 비동기 개체 감지 작업을 생성하고 실행합니다.

[자습서 시작하기 "](#)

- Amazon Comprehend를 사용하여 텍스트로 인사이트 분석

Amazon Comprehend를 사용하여 텍스트에서 인사이트를 분석하고 도출하는 방법을 알아봅니다.

[자습서 시작하기 "](#)

- Amazon Comprehend 요금

Amazon Comprehend 요금 및 예제에 대한 정보를 살펴봅니다.

[가이드 살펴보기 "](#)

Amazon Textract

- Amazon Textract 시작하기

형식이 지정된 텍스트와 함께 Amazon Textract를 사용하여 서로 가까운 단어와 단어 줄을 감지하고 관련 텍스트, 테이블, 키-값 페어, 선택 요소와 같은 항목에 대한 문서를 분석하는 방법을 알아봅니다.

[가이드 살펴보기 "](#)

- Amazon Textract를 사용하여 텍스트 및 정형 데이터 추출하기

Amazon Textract를 사용하여 문서에서 텍스트 및 구조화된 데이터를 추출하는 방법을 알아봅니다.

[자습서 시작하기 "](#)

- AWS 파워 아워: Machine Learning

이 에피소드에서는 Amazon Textract를 살펴보고,에서 시간을 보내고 AWS Management Console, 서비스 APIs를 최대한 활용하는 방법을 이해하는 데 도움이 되는 코드 샘플을 검토합니다.

[동영상 보기 "](#)

Amazon Translate

- 콘솔을 사용하여 Amazon Translate 시작하기

Amazon Translate를 시작하는 가장 쉬운 방법은 콘솔을 사용하여 텍스트를 번역하는 것입니다. 콘솔을 사용하여 최대 10,000자를 번역하는 방법을 알아봅니다.

[가이드 살펴보기 "](#)

- 클라우드의 언어 간 텍스트 번역

이 자습서 예제에서는 국제 여행 제조 회사의 일부로 고객이 현지 시장 언어인 프랑스어로 리뷰를 통해 제품에 대해 말하는 내용을 이해해야 합니다.

[자습서 시작하기 "](#)

- Amazon Translate 요금

12개월 동안 매월 2백만 문자를 제공하는 프리 티어를 포함한 Amazon Translate 요금을 살펴보세요.

[가이드 살펴보기 "](#)

Amazon Lex

- Amazon Lex V2 개발자 안내서

시작하기에 대한 정보, 작동 방식 및 Amazon Lex V2에 대한 요금 정보를 살펴봅니다.

[가이드 살펴보기 "](#)

- Amazon Lex 소개 Amazon Lex 대화형 서비스를 소개하고 봇을 생성하고 다른 채팅 서비스에 배포하는 방법을 보여주는 예제를 안내합니다.

[과정 수강 "\(로그인 필요\)](#)

- 대화형 경험에서 생성형 AI 탐색

대화 경험에서 생성형 AI의 사용을 살펴봅니다.

[블로그 읽기 "](#)

Amazon Polly

- Amazon Polly란 무엇입니까?

텍스트를 생생한 음성으로 변환하고 고객 참여와 접근성을 높이기 위한 애플리케이션을 개발하는 데 사용할 수 있는 클라우드 서비스에 대한 전체 개요를 살펴봅니다.

[가이드 살펴보기 "](#)

- Amazon Polly를 사용하여 말하는 텍스트를 강조 표시합니다.

책, 웹 사이트, 블로그 및 기타 디지털 경험에서 오디오에 시각적 기능을 추가하기 위해 텍스트를 강조 표시하는 방법을 소개합니다.

[블로그 읽기 "](#)

- Amazon Polly에서 동일한 TTS 음성 페르소나를 사용하여 여러 언어로 콘텐츠용 오디오 생성

NTTS(신경 Text-to-Speech)에 대해 설명하고 지원되는 언어로 다양한 개별 화자를 제공하는 사용 가능한 음성의 광범위한 포트폴리오가 어떻게 작동하는지 설명합니다.

[블로그 읽기 "](#)

Amazon Rekognition

- Amazon Rekognition이란 무엇인가요?

이 서비스를 사용하여 애플리케이션에 이미지 및 비디오 분석을 추가하는 방법을 알아봅니다.

[가이드 살펴보기 "](#)

- 실습 Rekognition: 자동 이미지 및 비디오 분석

코드 예제 및 키 포인트와 함께 스트리밍 비디오에서 얼굴 인식이 어떻게 작동하는지 자체 안내 속도로 알아봅니다.

[자습서 시작하기 "](#)

- Amazon Rekognition FAQs

Amazon Rekognition의 기본 사항과 이를 통해 딥 러닝을 개선하고 애플리케이션을 시각적으로 분석하는 방법을 알아봅니다.

[FAQs](#)

Amazon Transcribe

- Amazon Transcribe란 무엇입니까?

ML을 사용하여 오디오를 텍스트로 변환하는 AWS 자동 음성 인식 서비스를 살펴봅니다. 이 서비스를 독립 실행형 트랜스크립션으로 사용하거나 모든 애플리케이션에 speech-to-text 기능을 추가하는 방법을 알아봅니다.

[가이드 살펴보기 "](#)

- Amazon Transcribe 요금

사용자 지정 언어 모델 옵션 및 Amazon Transcribe 프리 티어를 포함한 AWS pay-as-you-go 트랜스크립션을 소개합니다.

[가이드 살펴보기 "](#)

- Amazon Transcribe를 사용하여 오디오 트랜스크립트 생성

Amazon Transcribe를 사용하여 요구 사항에 맞게 테스트하기 위한 실제 사용 사례 시나리오를 사용하여 녹음된 오디오 파일의 텍스트 트랜스크립트를 생성하는 방법을 알아봅니다.

[자습서 시작하기 "](#)

- Amazon Transcribe 스트리밍 앱 빌드

라이브 오디오를 실시간으로 녹음, 트랜스크립션 및 번역하는 앱을 빌드하고 결과를 직접 이메일로 보내는 방법을 알아봅니다.

[가이드 살펴보기 "](#)

두 번째 AI/ML AWS 서비스 경로 세트는 Amazon SageMaker AI 패밀리의 서비스를 시작하기 위한 심층적인 설명서, 실습 자습서 및 리소스를 제공합니다.

SageMaker AI

- Amazon SageMaker AI 작동 방식

기계 학습의 개요와 SageMaker AI의 작동 방식을 살펴봅니다.

[가이드 살펴보기 "](#)

- Amazon SageMaker AI 시작하기

Amazon SageMaker AI 도메인에 가입하여 Amazon SageMaker AI Studio 및 RStudio on SageMaker AI에 액세스할 수 있는 방법을 알아봅니다.

[가이드 살펴보기 "](#)

- Amazon SageMaker AI에서 Apache Spark 사용

데이터 사전 처리에 Apache Spark를 사용하고 모델 훈련 및 호스팅에 SageMaker AI를 사용하는 방법을 알아봅니다.

[가이드 살펴보기 "](#)

- Docker 컨테이너를 사용하여 모델 빌드

Amazon SageMaker AI가 빌드 및 런타임 작업에 Docker 컨테이너를 광범위하게 사용하는 방법을 알아봅니다. 기본 제공 알고리즘을 위해 사전 구축된 Docker 이미지와 훈련 및 추론에 사용되는 지원되는 딥 러닝 프레임워크를 배포하는 방법을 알아봅니다.

[가이드 살펴보기 "](#)

- 기계 학습 프레임워크 및 언어

Amazon SageMaker AI Python SDK를 사용하여 SageMaker AI를 시작하는 방법을 알아봅니다. Amazon SageMaker

[가이드 살펴보기 "](#)

SageMaker AI Autopilot

- 테이블 형식 데이터에 대한 Amazon SageMaker AI Autopilot 실험 생성

Amazon SageMaker AI Autopilot 실험을 생성하여 테이블 형식 데이터 세트에서 다양한 모델 후보를 탐색, 사전 처리 및 훈련하는 방법을 알아봅니다.

[가이드 살펴보기 "](#)

- 기계 학습 모델 자동 생성

Amazon SageMaker AI Autopilot을 사용하여 ML 모델을 자동으로 빌드, 훈련 및 튜닝하고 모델을 배포하여 예측하는 방법을 알아봅니다.

[자습서 시작하기 "](#)

- 다음 예제 노트북을 사용하여 Amazon SageMaker AI Autopilot을 사용한 모델링 살펴보기

다이렉트 마케팅, 고객 이탈 예측 및 자체 데이터 처리 코드를 Amazon SageMaker AI Autopilot으로 가져오는 방법에 대한 예제 노트북을 살펴보세요.

[가이드 살펴보기 "](#)

SageMaker AI Canvas

- Amazon SageMaker AI Canvas 사용 시작하기

SageMaker AI Canvas 사용을 시작하는 방법을 알아봅니다.

[가이드 살펴보기 "](#)

- 코드를 작성하지 않고 기계 학습 예측 생성

이 자습서에서는 Amazon SageMaker AI Canvas를 사용하여 한 줄의 코드를 작성하지 않고도 ML 모델을 구축하고 정확한 예측을 생성하는 방법을 설명합니다.

[자습서 시작하기 "](#)

- SageMaker AI Canvas 자세히 알아보기

SageMaker AI Canvas와 코드 없는 시각적 ML 기능을 자세히 살펴보세요.

[블로그 읽기 "](#)

- Amazon SageMaker AI Canvas를 사용하여 첫 번째 ML 모델 만들기

Amazon SageMaker AI Canvas를 사용하여 새로운 제품 및 서비스에 대한 이메일 캠페인을 기반으로 고객 보존을 평가하는 ML 모델을 생성하는 방법을 알아봅니다.

[랩 시작하기 "](#)

SageMaker AI Data Wrangler

- Amazon SageMaker AI Data Wrangler 시작하기

SageMaker AI Data Wrangler를 설정한 다음 기존 예제 데이터 세트를 사용하여 연습을 제공하는 방법을 알아봅니다.

[가이드 살펴보기 "](#)

- 최소한의 코드로 기계 학습을 위한 훈련 데이터 준비

Amazon SageMaker AI Data Wrangler를 사용하여 ML용 데이터를 준비하는 방법을 알아봅니다.

[자습서 시작하기 "](#)

- SageMaker AI Data Wrangler 심층 분석 워크숍

데이터 세트에 적절한 분석 유형을 적용하여 이상 및 문제를 감지하고, 파생 결과/통찰력을 사용하여 데이터 세트의 변환 과정에서 문제 해결 작업을 공식화하고, SageMaker AI Data Wrangler에서 제공하는 빠른 모델링 옵션을 사용하여 올바른 선택과 변환 시퀀스를 테스트하는 방법을 알아봅니다.

[워크숍 시작하기 "](#)

SageMaker AI Ground Truth

- Amazon Ground Truth 시작하기

콘솔을 사용하여 레이블 지정 작업을 생성하고, 퍼블릭 또는 프라이빗 작업 인력을 할당하고, 작업 인력에게 레이블 지정 작업을 보내는 방법을 알아봅니다. 레이블 지정 작업의 진행 상황을 모니터링하는 방법을 알아봅니다.

[가이드 살펴보기 "](#)

- Machine Learning을 위한 레이블 훈련 데이터

ML 모델의 훈련 데이터에 주석을 달기 위해 Amazon SageMaker AI Ground Truth에서 레이블 지정 작업을 설정하는 방법을 알아봅니다.

[자습서 시작하기 "](#)

- Amazon Ground Truth Plus 시작하기 Amazon SageMaker AI Ground Truth Plus 프로젝트를 시작하고, 레이블을 검토하고, SageMaker AI Ground Truth Plus 사전 조건을 충족하는 데 필요한 단계를 완료하는 방법을 알아봅니다.

[가이드 살펴보기 "](#)

- Amazon Ground Truth 시작하기 SageMaker AI Ground Truth 콘솔을 통해 몇 분 안에 데이터 레이블 지정을 시작하는 방법을 알아봅니다.

[동영상 보기 "](#)

- Amazon SageMaker AI Ground Truth Plus - 코드 또는 사내 리소스 없이 훈련 데이터 세트 생성
전문 인력을 사용하여 고품질 훈련 데이터 세트를 빠르게 제공하고 비용을 최대 40% 절감하는 터키 서비스인 Ground Truth Plus에 대해 알아봅니다.

[블로그 읽기 "](#)

SageMaker AI JumpStart

- SageMaker AI JumpStart로 기계 학습 시작하기

일반적인 사용 사례를 위한 인프라를 설정하는 SageMaker AI JumpStart 솔루션 템플릿과 SageMaker AI를 사용한 기계 학습을 위한 실행 가능한 예제 노트북을 살펴보세요.

[가이드 살펴보기 "](#)

- Amazon SageMaker AI JumpStart를 사용하여 기계 학습 프로젝트를 빠르게 시작하기

Amazon SageMaker AI JumpStart에서 제공하는 사전 훈련된 모델 및 사전 구축된 솔루션을 사용하여 ML 프로젝트를 빠르게 추적하는 방법을 알아봅니다. 그런 다음 Amazon SageMaker AI Studio 노트북을 통해 선택한 모델을 배포할 수 있습니다.

[자습서 시작하기 "](#)

- 이 침묵의 날 워크숍을 통해 Amazon SageMaker AI JumpStart 실습

Amazon SageMaker AI Data Wrangler, Autopilot 및 Jumpstart에서 찾은 로우 코드 ML 기능을 사용하여 더 빠르게 실험하고 매우 정확한 모델을 프로덕션에 제공하는 방법을 알아봅니다.

[워크숍 시작하기 "](#)

SageMaker AI Pipelines

- Amazon SageMaker AI 파이프라인 시작하기

SageMaker AI 작업을 관리하고 배포하는 end-to-end 워크플로를 생성하는 방법을 알아봅니다. SageMaker AI Pipelines은 SageMaker AI Python SDK 통합과 함께 제공되므로 Python 기반 인터페이스를 사용하여 파이프라인의 각 단계를 구축할 수 있습니다.

[가이드 살펴보기 "](#)

- 기계 학습 워크플로 자동화

Amazon SageMaker AI Pipelines, Amazon SageMaker AI 모델 레지스트리 및 Amazon SageMaker AI Amazon SageMaker Clarify를 사용하여 end-to-end 기계 학습(ML) 워크플로를 생성하고 자동화하는 방법을 알아봅니다.

[자습서 시작하기 "](#)

- Amazon SageMaker AI Pipelines을 사용하여 완전 자동화된 ML 워크플로를 생성하는 방법

모든 개발자와 데이터 과학자가 액세스할 수 있도록 설계된 세계 최초의 ML CI/CD 서비스인 Amazon SageMaker AI Pipelines에 대해 알아봅니다. SageMaker AI 파이프라인은 ML에 CI/CD 파이프라인을 제공하여 필요한 코딩 시간을 줄입니다.

[동영상 보기 "](#)

SageMaker AI Studio

- 로컬에서 기계 학습 모델 구축 및 훈련

Amazon SageMaker AI Studio 노트북 내에서 로컬로 ML 모델을 빌드하고 훈련하는 방법을 알아봅니다.

[자습서 시작하기 "](#)

- SageMaker AI Studio와 EMR 워크숍 통합

대규모 분산 처리를 활용하여 데이터를 준비한 다음 기계 학습 모델을 훈련하는 방법을 알아봅니다.

[워크숍 시작하기 "](#)

세 번째 AI/ML AWS 서비스 경로 세트는 , 및 Amazon Titan을 시작하기 위한 심층적인 설명서, 실습 자습서 AWS Trainium AWS Inferentia 및 리소스를 제공합니다.

AWS Trainium

- AWS Trainium 및 Amazon EKS를 사용하여 분산 훈련 확장

클라우드에서 딥 러닝 모델을 훈련하기 위해 비용 효율적이고 대규모로 확장 가능한 고성능 플랫폼을 제공하도록 최적화된 목적별 ML 액셀러레이터 AWS Trainium인 로 구동되는 Amazon EC2 Trn1 인스턴스의 일반 가용성을 활용하는 방법을 알아봅니다.

[블로그 읽기 "](#)

- 개요 AWS Trainium

100B 개 이상의 파라미터 모델의 딥 러닝 훈련을 위해 AWS 특별히 구축된 AWS Trainium 2세대 기계 학습(ML) 액셀러레이터에 대해 알아봅니다. 각 Amazon Elastic Compute Cloud(EC2) Trn1 인스턴스는 최대 16개의 AWS Trainium 가속기를 배포하여 클라우드에서 딥 러닝(DL) 훈련을 위한 고성능 저비용 솔루션을 제공합니다.

[가이드 살펴보기 "](#)

- Recommended Trainium Instances

딥 러닝 모델 추론 워크로드에 고성능 및 비용 효율성을 제공하도록 AWS Trainium 인스턴스를 설계하는 방법을 알아봅니다.

[가이드 살펴보기 "](#)

AWS Inferentia

- 개요 AWS Inferentia

가 딥 러닝(DL) 추론 애플리케이션에 최저 비용으로 고성능 AWS 을 제공하도록 액셀러레이터를 설계하는 방법을 이해합니다.

[가이드 살펴보기 "](#)

- AWS Inferentia 4배 더 높은 처리량과 AWS Inferentia 10배 더 짧은 지연 시간을 제공하여 2개 빌드1

에 최적화된 what AWS Inferentia 2를 이해하고 LLMs 및 생성형 AI 추론 비용을 낮추면서 더 높은 성능을 제공하도록 처음부터 어떻게 설계되었는지 살펴봅니다.

[블로그 읽기 "](#)

- 를 사용한 기계 학습 추론 AWS Inferentia

Amazon EC2 Inf1 인스턴스를 실행하는 노드가 있는 Amazon EKS 클러스터를 생성하고 (선택 사항) 샘플 애플리케이션을 배포하는 방법을 알아봅니다. Inf1 Amazon EC2 Inf1 인스턴스는 클라우드에서 고성능 및 최저 비용 추론을 제공하기 AWS 위해에서 사용자 지정한 AWS Inferentia 칩으로 구동됩니다.

[가이드 살펴보기 "](#)

Amazon Titan

- Amazon Titan 개요

Amazon Titan FMs 대규모 데이터 세트에서 사전 훈련되어 강력한 범용 모델이 되는 방법을 알아봅니다. 대량의 데이터에 주석을 달지 않고 특정 작업에 대한 자체 데이터로 사용자 지정하는 방법을 알아봅니다.

[가이드 살펴보기 "](#)

탐색

- 아키텍처 다이어그램

이러한 참조 아키텍처 다이어그램은 사용 중인 AWS AI 및 ML 서비스의 예를 보여줍니다.

[아키텍처 다이어그램 살펴보기 "](#)

- 백서

시작에 도움이 되는 백서를 살펴보고 AI/ML 서비스를 선택하고 사용하는 모범 사례를 알아봅니다.

[백서 살펴보기 "](#)

- AWS 솔루션

AI 및 ML 서비스의 일반적인 사용 사례에 대한 검증된 솔루션 및 아키텍처 지침을 살펴보세요.

[솔루션 살펴보기 "](#)

리소스

파운데이션 모델

지원되는 파운데이션 모델은 다음과 같습니다.

- [엔트로픽 클로드](#)
- [Cohere 명령 및 임베드](#)
- [AI21 Labs Jurassic](#)
- [메타 라마](#)
- [Mistral AI](#)
- [안정적인 확산 XL](#)
- [Amazon Titan](#)

Amazon Bedrock을 사용하면 다양한 파운데이션 모델을 실험하고 데이터로 비공개로 사용자 지정할 수 있습니다.

사용 사례 또는 업계별 서비스

- [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](#)

관련 블로그 게시물

- [중요한 새로운 기능을 통해 Amazon Bedrock을 사용하여 생성형 AI 애플리케이션을 더 쉽게 구축 및 확장하고 놀라운 결과를 얻을 수 있습니다.](#)
- [AWS Inferentia 및 Amazon SageMaker AI JumpStart에서 Llama 3 모델 배포를 위한 최저 비용 AWS Trainium 제공](#)
- [Amazon SageMaker AI에서 비즈니스를 위한 맞춤형 보상 모델로 고객 만족도 혁신](#)
- [Personalize, 지연 시간이 짧은 더 큰 항목 카탈로그를 지원하는 새로운 레시피 출시](#)

문서 이력

다음 표에서는 이 결정 가이드의 중요한 변경 사항에 대해 설명합니다. 이 가이드의 업데이트에 대한 알림을 받으려면 RSS 피드를 구독하면 됩니다.

변경 사항	설명	날짜
마이너 업데이트	Amazon Q 및 Amazon의 최신 AI 및 ML 스택에 대한 콘텐츠가 업데이트되었습니다.	2024년 5월 3일
최초 릴리스	결정 가이드의 최초 릴리스.	2023년 7월 24일

기계 번역으로 제공되는 번역입니다. 제공된 번역과 원본 영어의 내용이 상충하는 경우에는 영어 버전이 우선합니다.