



AWS Leitfaden zur Entscheidungsfindung

Auswahl eines Dienstes AWS für maschinelles Lernen



Auswahl eines Dienstes AWS für maschinelles Lernen: AWS Leitfaden zur Entscheidungsfindung

Copyright © 2026 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Die Marken und Handelsmarken von Amazon dürfen nicht in einer Weise in Verbindung mit nicht von Amazon stammenden Produkten oder Services verwendet werden, die geeignet ist, die Kunden zu verwirren oder Amazon in einer Weise herabzusetzen oder zu diskreditieren. Alle anderen Marken, die nicht Eigentum von Amazon sind, sind Eigentum ihrer jeweiligen Inhaber, die mit Amazon verbunden oder nicht verbunden oder von Amazon gesponsert oder nicht gesponsert sein können.

Table of Contents

Leitfaden zur Entscheidungsfindung	1
Einführung	1
Verstehe	2
Überlegen	4
Klicken Sie auf	7
Verwenden Sie	10
Erkunden	23
Ressourcen	24
Dokumentverlauf	26
.....	xxvii

Auswahl eines Dienstes AWS für maschinelles Lernen

Wählen Sie die richtigen ML-Services und Frameworks zur Unterstützung Ihrer Arbeit

Zweck	Finden Sie heraus, welche AWS ML-Services am besten zu Ihren Anforderungen passen.
Letzte Aktualisierung	3. Mai 2024
Abgedeckte Dienstleistungen	<ul style="list-style-type: none">• Amazon Augmented AI• Amazon CodeGuru• Amazon Comprehend• DevOpsAmazon-Guru• Amazon Forecast• Amazon Kendra• Amazon Lex• Amazon Personalize• Amazon Polly• Amazon Rekognition• Amazon SageMaker KI• Amazon Textract• Amazon Transcribe• Amazon Translate

Einführung

Im Grunde ist maschinelles Lernen (ML) darauf ausgelegt, digitale Tools und Dienste bereitzustellen, um aus Daten zu lernen, Muster zu identifizieren, Vorhersagen zu treffen und dann auf diese Vorhersagen zu reagieren. Fast alle Systeme der künstlichen Intelligenz (KI) werden heute mithilfe von ML erstellt. ML verwendet große Datenmengen, um Entscheidungslogik zu erstellen und zu validieren. Diese Entscheidungslogik bildet die Grundlage des KI-Modells.

Zu den Szenarien, in denen Dienste für AWS maschinelles Lernen eingesetzt werden können, gehören:

- **Spezifische Anwendungsfälle** — Dienste für AWS maschinelles Lernen können Ihre KI-gestützten Anwendungsfälle mit einer breiten Palette von vorgefertigten Algorithmen, Modellen und Lösungen für gängige Anwendungsfälle und Branchen unterstützen. Sie haben die Wahl zwischen 23 vortrainierten Services, darunter Amazon Personalize, Amazon Kendra und Amazon Monitron.
- **Anpassung und Skalierung von maschinellem Lernen** — Amazon SageMaker AI wurde entwickelt, um Sie beim Erstellen, Trainieren und Bereitstellen von ML-Modellen für jeden Anwendungsfall zu unterstützen. Sie können Ihre eigenen erstellen oder AWS über Amazon SageMaker AI und Amazon Bedrock auf Open-Source-Basismodelle zugreifen.
- **Zugriff auf spezielle Infrastrukturen** — Verwenden Sie die ML-Frameworks und die Infrastruktur von, AWS wenn Sie noch mehr Flexibilität und Kontrolle über Ihre Workflows für maschinelles Lernen benötigen und bereit sind, die zugrunde liegende Infrastruktur und Ressourcen selbst zu verwalten.

Dieser Entscheidungsleitfaden hilft Ihnen dabei, die richtigen Fragen zu stellen, Ihre Kriterien und Ihr Geschäftsproblem zu bewerten und festzustellen, welche Services am besten zu Ihren Anforderungen passen.

Verstehen

Da Unternehmen weiterhin KI- und ML-Technologien einsetzen, ist es eine ständige Herausforderung, AWS ML-Services zu verstehen und zwischen ihnen auszuwählen.

AWS bietet eine Reihe von ML-Services, die Unternehmen dabei unterstützen sollen, ML-Modelle schneller und einfacher zu erstellen, zu trainieren und einzusetzen. Diese Dienste können zur Lösung einer Vielzahl von Geschäftsproblemen eingesetzt werden, z. B. zur Vorhersage von Kundenabwanderungen, Betrugserkennung sowie Bild- und Spracherkennung.

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)

Bevor wir uns eingehender mit AWS ML-Diensten befassen, schauen wir uns die Beziehung zwischen KI und ML an.

- Auf einer höheren Ebene ist künstliche Intelligenz eine Möglichkeit, jedes System zu beschreiben, das Aufgaben replizieren kann, für die zuvor menschliche Intelligenz erforderlich war. Bei den meisten KI-Anwendungsfällen geht es um ein probabilistisches Ergebnis — das Treffen einer Vorhersage oder Entscheidung mit einem hohen Maß an Sicherheit, ähnlich dem menschlichen Urteilsvermögen.
- Fast alle KI-Systeme werden heute mithilfe von maschinellem Lernen erstellt. ML verwendet große Datenmengen, um eine Entscheidungslogik zu erstellen und zu validieren, die als Modell bezeichnet wird.
- Klassifikation KI ist eine Untergruppe von ML, die Muster erkennt, um etwas zu identifizieren. Prädiktive KI ist eine Untergruppe von ML, die future Trends auf der Grundlage statistischer Muster und historischer Daten vorhersagt.
- Schließlich ist generative KI eine Untergruppe von Deep Learning, mit der neue Inhalte und Ideen wie Konversationen, Geschichten, Bilder, Videos und Musik entstehen können. Generative KI basiert auf sehr großen Modellen, die anhand riesiger Datenmengen vorab trainiert wurden. Sie werden als Foundation-Modelle oder bezeichnet. FMs [Amazon Bedrock](#) ist ein vollständig

verwalteter Service, der eine Auswahl an leistungsstarken Anwendungen FMs für die Erstellung und Skalierung generativer KI-Anwendungen bietet. [Amazon Q Developer](#) und [Amazon Q Business](#) sind Assistenten mit generativer KI für bestimmte Anwendungsfälle.

Dieser Leitfaden befasst sich in erster Linie mit Dienstleistungen in den Kategorien Classification AI und Predictive AI Machine Learning.

Darüber hinaus AWS bietet es spezialisierte, beschleunigte Hardware für leistungsstarkes ML-Training und Inferenz.

- [Amazon EC2 P5-Instances](#) sind mit NVIDIA H100 Tensor Core ausgestattet GPUs, die sich sowohl für Schulungs- als auch für Inferenzaufgaben beim maschinellen Lernen eignen. [Amazon EC2 G5-Instances](#) verfügen über bis zu 8 NVIDIA A10G Tensor Core GPUs - und AMD EPYC-Prozessoren der zweiten Generation für eine Vielzahl von grafikintensiven Anwendungsfällen und maschinellem Lernen.
- [AWS Trainium](#) ist der ML-Beschleuniger der zweiten Generation, der speziell für das Deep-Learning-Training (AWS DL) von Modellen mit mehr als 100 Parametern entwickelt wurde.
- [AWS Inferentia 2-basierte Amazon EC2 Inf2-Instances](#) sind darauf ausgelegt, hohe Leistung zu den niedrigsten Kosten in Amazon EC2 für Ihre DL- und generativen KI-Inferenzanwendungen zu bieten.

Überlegen Sie

Bei der Lösung eines Geschäftsproblems mit AWS ML-Services kann die Berücksichtigung mehrerer wichtiger Kriterien zum Erfolg beitragen. Im folgenden Abschnitt werden einige der wichtigsten Kriterien beschrieben, die bei der Auswahl eines ML-Services zu berücksichtigen sind.

Problem definition

Definition des Problems

Der erste Schritt im ML-Lebenszyklus besteht darin, das Geschäftsproblem zu formulieren. Für die Auswahl des richtigen AWS ML-Services ist es wichtig, das Problem zu verstehen, das Sie zu lösen versuchen, da verschiedene Dienste darauf ausgelegt sind, unterschiedliche Probleme zu lösen. Es ist auch wichtig festzustellen, ob ML für Ihr Geschäftsproblem am besten geeignet ist.

Sobald Sie festgestellt haben, dass ML am besten geeignet ist, können Sie zunächst aus einer Reihe von speziell entwickelten AWS KI-Services (in Bereichen wie Sprache, Vision und Dokumente) wählen.

Amazon SageMaker AI bietet eine vollständig verwaltete Infrastruktur, wenn Sie Ihre eigenen Modelle erstellen und trainieren müssen. AWS bietet eine Reihe fortschrittlicher ML-Frameworks und Infrastrukturoptionen für die Fälle, in denen Sie hochgradig maßgeschneiderte und spezialisierte ML-Modelle benötigen. AWS bietet außerdem eine breite Palette beliebter Basismodelle für die Entwicklung neuer Anwendungen mit generativer KI.

ML algorithm

ML-Algorithmus

Die Wahl des ML-Algorithmus für das Geschäftsproblem, das Sie lösen möchten, hängt von der Art der Daten ab, mit denen Sie arbeiten, sowie von den gewünschten Ergebnissen. Die folgenden Informationen beschreiben, wie Sie mit jeder der wichtigsten AWS KI/ML-Servicekategorien mit ihren Algorithmen arbeiten können:

- **Spezialisierte KI-Dienste:** Diese Dienste bieten nur begrenzte Möglichkeiten zur Anpassung des ML-Algorithmus, da es sich um vortrainierte Modelle handelt, die für bestimmte Aufgaben optimiert sind. In der Regel können Sie die Eingabedaten und einige Parameter anpassen, haben jedoch keinen Zugriff auf die zugrunde liegenden ML-Modelle oder die Möglichkeit, Ihre eigenen Modelle zu erstellen.
- **Amazon SageMaker AI:** Dieser Service bietet die größte Flexibilität und Kontrolle über den ML-Algorithmus. Sie können SageMaker KI verwenden, um benutzerdefinierte Modelle mit Ihren eigenen Algorithmen und Frameworks zu erstellen, oder Sie können vorgefertigte Modelle und Algorithmen verwenden, die von AWS bereitgestellt werden. Dies ermöglicht ein hohes Maß an Anpassung und Kontrolle über den ML-Prozess.
- **ML-Frameworks und Infrastruktur auf niedrigerer Ebene:** Diese Dienste bieten die größte Flexibilität und Kontrolle über den ML-Algorithmus. Sie können diese Dienste verwenden, um mithilfe ihrer eigenen Algorithmen und Frameworks hochgradig maßgeschneiderte ML-Modelle zu erstellen. Die Nutzung dieser Dienste erfordert jedoch umfangreiche ML-Kenntnisse und ist möglicherweise nicht für alle Anwendungsfälle durchführbar.

Security

Sicherheit

Wenn Sie einen privaten Endpunkt in Ihrer VPC benötigen, variieren Ihre Optionen je nach der Ebene der AWS ML-Services, die Sie verwenden. Dazu zählen:

- **Spezialisierte KI-Dienste:** Die meisten spezialisierten KI-Dienste unterstützen derzeit keine privaten Endpunkte in VPCs. Auf Amazon Rekognition Custom Labels und Amazon Comprehend Custom kann jedoch über VPC-Endpunkte zugegriffen werden.
- **Kern-KI-Dienste:** Amazon Translate, Amazon Transcribe und Amazon Comprehend unterstützen alle VPC-Endpunkte.
- **Amazon SageMaker AI:** SageMaker AI bietet integrierte Unterstützung für VPC-Endpunkte, sodass Sie deren trainierte Modelle als Endpunkt bereitstellen können, auf den nur von deren VPC aus zugegriffen werden kann.
- **ML-Frameworks und Infrastruktur auf niedrigerer Ebene:** Sie können Ihre Modelle auf EC2 Amazon-Instances oder in Containern innerhalb Ihrer VPC bereitstellen und so die vollständige Kontrolle über die Netzwerkkonfiguration haben.

Latency

Latency

Höhere KI-Dienste wie Amazon Rekognition und Amazon Transcribe sind für eine Vielzahl von Anwendungsfällen konzipiert und bieten eine hohe Geschwindigkeit. Sie erfüllen jedoch möglicherweise bestimmte Latenzanforderungen nicht.

Wenn Sie ML-Frameworks und Infrastrukturen auf niedrigerer Ebene verwenden, empfehlen wir die Nutzung von Amazon AI. SageMaker. Diese Option ist aufgrund des vollständig verwalteten Services und der optimierten Bereitstellungsoptionen im Allgemeinen schneller als die Erstellung benutzerdefinierter Modelle. Ein hochoptimiertes kundenspezifisches Modell mag zwar besser abschneiden als SageMaker KI, erfordert jedoch umfangreiche Fachkenntnisse und Ressourcen, um es zu entwickeln.

Accuracy

Genauigkeit

Die Genauigkeit von AWS ML-Diensten hängt vom jeweiligen Anwendungsfall und dem Grad der erforderlichen Anpassung ab. Höhere KI-Dienste wie Amazon Rekognition basieren auf vorab trainierten Modellen, die für bestimmte Aufgaben optimiert wurden und in vielen Anwendungsfällen eine hohe Genauigkeit bieten.

In einigen Fällen können Sie sich für Amazon SageMaker AI entscheiden, das eine flexiblere und anpassbarere Plattform für die Erstellung und das Training benutzerdefinierter ML-Modelle bietet. Wenn Sie Ihre eigenen Modelle erstellen, können Sie möglicherweise eine noch höhere Genauigkeit erzielen, als dies mit vortrainierten Modellen möglich ist.

Sie können sich auch dafür entscheiden, ML-Frameworks und -Infrastrukturen wie TensorFlow Apache zu verwenden MXNet, um hochgradig maßgeschneiderte Modelle zu erstellen, die die höchstmögliche Genauigkeit für Ihren spezifischen Anwendungsfall bieten.

AWS and responsible AI

AWS und verantwortungsvolle KI

AWS erstellt Grundmodelle (FMs), bei denen eine verantwortungsvolle KI in jeder Phase des Entwicklungsprozesses berücksichtigt wird. Während des Entwurfs, der Entwicklung, der Bereitstellung und des Betriebs berücksichtigen wir eine Reihe von Faktoren, darunter:

1. Genauigkeit (wie sehr eine Zusammenfassung mit dem zugrunde liegenden Dokument übereinstimmt; ob eine Biographie sachlich korrekt ist)
2. Fairness (ob die Ergebnisse demografische Gruppen ähnlich behandeln)
3. Überlegungen zum geistigen Eigentum und zum Urheberrecht
4. Angemessene Verwendung (Herausfiltern von Benutzeranfragen nach Rechtsberatung, medizinischen Diagnosen oder illegalen Aktivitäten)
5. Toxizität (Hassreden, Obszönitäten und Beleidigungen)
6. Datenschutz (Schutz personenbezogener Daten und Kundenaufforderungen)

AWS integriert Lösungen zur Lösung dieser Probleme in die Prozesse zur Erfassung von Trainingsdaten, in die Prozesse FMs selbst und in die Technologie, die zur Vorverarbeitung von Benutzeraufforderungen und zur Nachverarbeitung von Ergebnissen verwendet wird.

Klicken Sie auf

Nachdem Sie nun die Kriterien kennen, anhand derer Sie Ihre ML-Serviceoptionen bewerten werden, können Sie entscheiden, welcher AWS ML-Service für Ihre Unternehmensanforderungen am besten geeignet ist. In der folgenden Tabelle wird hervorgehoben, welche ML-Services für welche Umstände optimiert sind. Verwenden Sie es, um den AWS ML-Service zu ermitteln, der für Ihren Anwendungsfall am besten geeignet ist.

Kategorien	Wann würdest du es benutzen?	Wofür ist es optimiert?	Verwandte AI/ML Dienste oder Umgebungen
<p>Spezifische Anwendungsfälle</p> <p>Diese Dienste der künstlichen Intelligenz sollen spezifische Bedürfnisse erfüllen. Dazu gehören Personalisierung, Prognose, Erkennung von Anomalien, Sprachtranskription und andere. Da sie als Dienste bereitgestellt werden, können sie in Anwendungen eingebettet werden, ohne dass ML-Kenntnisse erforderlich sind.</p>	<p>Nutzen Sie die von bereitgestellten KI-Services, AWS wenn Sie spezielle, vorgefertigte Funktionen in Ihre Anwendungen integrieren möchten, ohne dass umfangreiche Anpassungen oder Fachwissen im Bereich maschinelles Lernen erforderlich sind. Diese Dienste sind so konzipiert, dass sie einfach zu bedienen sind und nicht viel Codierung oder Konfiguration erfordern.</p>	<p>Diese Dienste sind so konzipiert, dass sie einfach zu bedienen sind und nicht viel Fachwissen in den Bereichen Codierung, Konfiguration oder maschinelles Lernen erfordern.</p>	<p>Amazon Augmented AI</p> <p>Amazon CodeGuru</p> <p>Amazon Comprehend</p> <p>Amazon Comprehend Medical</p> <p>DevOpsAmazon-Guru</p> <p>Amazon Forecast</p> <p>Amazon Kendra</p> <p>Amazon Lex</p> <p>Amazon Personalize</p> <p>Amazon Polly</p> <p>Amazon Rekognition</p> <p>Amazon Textract</p> <p>Amazon Transcribe</p> <p>Amazon Translate</p>
<p>ML-Dienste</p> <p>Diese Dienste können verwendet werden, um maßgeschneiderte Modelle oder Workflows für</p>	<p>Verwenden Sie diese Dienste, wenn Sie individuellere Modelle oder Workflows für maschinelles Lernen benötigen, die über die vorgefertigten</p>	<p>Diese Dienste sind für die Erstellung und das Training benutzerdefinierter Modelle für maschinelles Lernen, groß angelegte Schulungen auf</p>	<p>Amazon SageMaker KI</p> <p>Amazon SageMaker KI JumpStart</p> <p>SageMaker KI-Studio</p>

Kategorien	Wann würdest du es benutzen?	Wofür ist es optimiert?	Verwandte AI/ML Dienste oder Umgebungen
<p>maschinelles Lernen zu entwickeln, die über die vorgefertigten Funktionen der KI-Kerndienste hinausgehen.</p>	<p>igten Funktionen der KI-Kerndienste hinausgehen.</p>	<p>mehreren Instanzen oder GPU-Clustern, mehr Kontrolle über die Bereitstellung von Modellen für maschinelles Lernen, Inferenzen in Echtzeit und die Erstellung end-to-end von Workflows optimiert.</p>	<p>SageMaker KI-Leinwand</p> <p>SageMaker Labor im KI-Studio</p> <p>SageMaker KI Ground Truth</p> <p>PyTorch auf AWS</p> <p>Apache MxNet</p> <p>Hugging Face</p> <p>TensorFlow auf AWS</p>
<p>Infrastruktur</p> <p>Um maschinelles Lernen in der Produktion einzusetzen, benötigen Sie eine kostengünstige Infrastruktur, die Amazon mit AWS eingebautem Silizium ermöglicht.</p>	<p>Verwenden Sie diese Option, wenn Sie die niedrigsten Kosten für Trainingsmodelle erzielen möchten und Inferenz in der Cloud ausführen müssen.</p>	<p>Optimiert für die Unterstützung des kostengünstigen Einsatzes von maschinellem Lernen.</p>	<p>AWS Trainium</p> <p>AWS Inferentia und Inferentia2</p> <p>Amazon SageMaker KI HyperPod</p>

Kategorien	Wann würdest du es benutzen?	Wofür ist es optimiert?	Verwandte AI/ML Dienste oder Umgebungen
<p>Tools und zugehörige Dienste</p> <p>Diese Tools und die zugehörigen Dienste sollen Ihnen helfen, den Einsatz von maschinellem Lernen zu vereinfachen.</p>	<p>Diese Services und Tools sollen Ihnen helfen, Deep Learning in der Cloud zu beschleunigen, indem sie Amazon-Maschinen-Images, Docker-Images und Entitätsauflösung bereitstellen.</p>	<p>Optimiert, um Ihnen zu helfen, Deep Learning in der Cloud zu beschleunigen.</p>	<p>AWS Deep Learning AMIs S</p> <p>AWS Deep Learning Containers</p> <p>AWS Auflösung von Entitäten</p>

Verwenden Sie

Nachdem Sie nun ein klares Verständnis der Kriterien haben, die Sie bei der Auswahl eines AWS ML-Services anwenden müssen, können Sie auswählen, welche AWS KI/ML-Services für Ihre Geschäftsanforderungen optimiert sind.

Um herauszufinden, wie Sie die Dienste, die Sie ausgewählt haben, nutzen und mehr über sie erfahren können, haben wir drei Arten von Pfaden bereitgestellt, um zu untersuchen, wie die einzelnen Dienste funktionieren. Die erste Reihe von Pfaden bietet ausführliche Dokumentation, praktische Tutorials und Ressourcen für den Einstieg in Amazon Comprehend, Amazon Textract, Amazon Translate, Amazon Lex, Amazon Polly, Amazon Rekognition und Amazon Transcribe.

Amazon Comprehend

- Erste Schritte mit Amazon Comprehend

Verwenden Sie die Amazon Comprehend Comprehend-Konsole, um einen asynchronen Entitätserkennungsauftrag zu erstellen und auszuführen.

[Beginnen Sie mit dem Tutorial](#)

- Analysieren Sie Erkenntnisse in Text mit Amazon Comprehend

Erfahren Sie, wie Sie Amazon Comprehend verwenden, um Text zu analysieren und Erkenntnisse daraus abzuleiten.

[Beginnen Sie mit dem Tutorial“](#)

- Amazon Comprehend — Preisgestaltung

Informieren Sie sich über die Preise und Beispiele von Amazon Comprehend.

[Erkunden Sie den Leitfaden“](#)

Amazon Textract

- Erste Schritte mit Amazon Textract

Erfahren Sie, wie Amazon Textract mit formatiertem Text verwendet werden kann, um Wörter und Wortzeilen zu erkennen, die nahe beieinander liegen, und ein Dokument auf Elemente wie verwandten Text, Tabellen, Schlüssel-Wert-Paare und Auswahrelemente zu analysieren.

[Erkunden Sie den Leitfaden“](#)

- Extrahieren Sie Text und strukturierte Daten mit Amazon Textract

Erfahren Sie, wie Sie mit Amazon Textract Text und strukturierte Daten aus einem Dokument extrahieren.

[Beginnen Sie mit dem Tutorial“](#)

- AWS Power Hour: Machine Learning

Tauchen Sie in dieser Episode in Amazon Textract ein, verbringen Sie Zeit damit und sehen Sie sich Codebeispiele an AWS-Managementkonsole, die Ihnen helfen, zu verstehen, wie Sie den Service APIs optimal nutzen können.

[Sehen Sie sich das Video an“](#)

Amazon Translate

- Erste Schritte mit Amazon Translate über die Konsole

Der einfachste Weg, mit Amazon Translate zu beginnen, besteht darin, die Konsole zu verwenden, um Text zu übersetzen. Erfahren Sie, wie Sie mit der Konsole bis zu 10.000 Zeichen übersetzen können.

[Erkunden Sie den Leitfaden“](#)

- Translate Sie Text zwischen Sprachen in der Cloud

In diesem Tutorial-Beispiel müssen Sie als Teil eines internationalen Kofferherstellers verstehen, was Kunden in Bewertungen in der lokalen Marktsprache Französisch über Ihr Produkt sagen.

[Fangen Sie mit dem Tutorial an“](#)

- Preise für Amazon Translate

Informieren Sie sich über die Preise von Amazon Translate, einschließlich des kostenlosen Kontingents, das 12 Monate lang 2 Millionen Zeichen pro Monat bereitstellt.

[Erkunden Sie den Leitfaden“](#)

Amazon Lex

- Amazon Lex V2 Entwicklerhandbuch

Informieren Sie sich über die ersten Schritte, die Funktionsweise und Preisinformationen für Amazon Lex V2.

[Erkunden Sie den Leitfaden“](#)

- Einführung in Amazon Lex Wir stellen Ihnen den Amazon Lex Conversational Service vor und führen Sie durch Beispiele, die Ihnen zeigen, wie Sie einen Bot erstellen und ihn für verschiedene Chat-Dienste einsetzen.

[Nehmen Sie am Kurs teil“](#) (Anmeldung erforderlich)

- Erkundung generativer KI in Konversationserlebnissen

Erkunden Sie den Einsatz generativer KI in Konversationserlebnissen.

[Lesen Sie den Blog](#)

Amazon Polly

- Was ist Amazon Polly?

Verschaffen Sie sich einen vollständigen Überblick über den Cloud-Service, der Text in lebensechte Sprache umwandelt und zur Entwicklung von Anwendungen verwendet werden kann, mit denen Sie Ihre Kundenbindung und Barrierefreiheit verbessern können.

[Erkunden Sie den Leitfaden](#)

- Markieren Sie Text, während er mit Amazon Polly gesprochen wird

Wir stellen Ihnen Methoden vor, mit denen Sie Text beim Sprechen hervorheben können, um Audio in Büchern, Websites, Blogs und anderen digitalen Erlebnissen um visuelle Funktionen zu erweitern.

[Lesen Sie den Blog](#)

- Erstellen Sie Audio für Inhalte in mehreren Sprachen mit derselben TTS-Stimme in Amazon Polly

Wir erklären Neural Text-to-Speech (NTTS) und besprechen, wie ein breites Portfolio verfügbarer Stimmen, das eine Reihe von unterschiedlichen Sprechern in unterstützten Sprachen bereitstellt, für Sie funktionieren kann.

[Lesen Sie den Blog](#)

Amazon Rekognition

- Was ist Amazon Rekognition?

Erfahren Sie, wie Sie diesen Service nutzen können, um Ihren Anwendungen Bild- und Videoanalysen hinzuzufügen.

[Erkunden Sie den Leitfaden“](#)

- Rekognition zum Anfassen: Automatisierte Bild- und Videoanalyse

Erfahren Sie in selbstgesteuertem Tempo, wie Gesichtserkennung mit Streaming-Videos funktioniert, zusammen mit Codebeispielen und wichtigen Punkten.

[Fangen Sie mit dem Tutorial an“](#)

- Amazon Rekognition FAQs

Lernen Sie die Grundlagen von Amazon Rekognition kennen und erfahren Sie, wie es Ihnen helfen kann, Ihr Deep Learning zu verbessern und Ihre Anwendungen visuell zu analysieren.

[Lesen Sie das“ FAQs](#)

Amazon Transcribe

- Was ist Amazon Transcribe?

Entdecken Sie den AWS automatischen Spracherkennungsdienst, der ML verwendet, um Audio in Text umzuwandeln. Erfahren Sie, wie Sie diesen Dienst als eigenständige Transkription verwenden oder einer beliebigen Anwendung speech-to-text Funktionen hinzufügen können.

[Erkunden Sie den Leitfaden“](#)

- Amazon Transcribe — Preise

Wir stellen Ihnen die AWS pay-as-you-go Transkription vor, einschließlich benutzerdefinierter Sprachmodelloptionen und des kostenlosen Kontingents für Amazon Transcribe.

[Erkunden Sie den Leitfaden“](#)

- Erstellen Sie ein Audiotranskript mit Amazon Transcribe

Erfahren Sie, wie Sie Amazon Transcribe verwenden, um ein Texttranskript von aufgezeichneten Audiodateien zu erstellen, indem Sie ein reales Anwendungsszenario verwenden, um es anhand Ihrer Anforderungen zu testen.

[Beginnen Sie mit dem Tutorial“](#)

- Erstellen Sie eine Amazon Transcribe Transcribe-Streaming-App

Erfahren Sie, wie Sie eine App erstellen, mit der Sie Live-Audio in Echtzeit aufnehmen, transkribieren und übersetzen können. Die Ergebnisse werden Ihnen direkt per E-Mail zugeschickt.

[Erkunden Sie den Leitfaden“](#)

Die zweite Gruppe von AWS-Service KI/ML-Pfaden bietet ausführliche Dokumentation, praktische Tutorials und Ressourcen für den Einstieg in die Dienste der Amazon SageMaker AI-Familie.

SageMaker AI

- So funktioniert Amazon SageMaker AI

Erkunden Sie den Überblick über maschinelles Lernen und die Funktionsweise von SageMaker KI.

[Erkunden Sie den Leitfaden“](#)

- Erste Schritte mit Amazon SageMaker AI

Erfahren Sie, wie Sie einer Amazon SageMaker AI-Domain beitreten und so Zugriff RStudio auf Amazon SageMaker AI Studio und SageMaker KI erhalten.

[Erkunden Sie den Leitfaden“](#)

- Verwenden Sie Apache Spark mit Amazon SageMaker AI

Erfahren Sie, wie Sie Apache Spark für die Vorverarbeitung von Daten und SageMaker KI für Modelltraining und Hosting verwenden.

[Erkunden Sie den Leitfaden“](#)

- Verwenden Sie Docker-Container, um Modelle zu erstellen

Erfahren Sie, wie Amazon SageMaker AI Docker-Container in großem Umfang für Build- und Runtime-Aufgaben nutzt. Erfahren Sie, wie Sie die vorgefertigten Docker-Images für die integrierten Algorithmen und die unterstützten Deep-Learning-Frameworks einsetzen, die für Training und Inferenz verwendet werden.

[Erkunden Sie den Leitfaden“](#)

- Frameworks und Sprachen für maschinelles Lernen

Erfahren Sie, wie Sie mithilfe des Amazon SageMaker AI Python SDK mit SageMaker KI beginnen können.

[Erkunden Sie den Leitfaden“](#)

SageMaker AI Autopilot

- Erstellen Sie ein Amazon SageMaker AI Autopilot-Experiment für tabellarische Daten

Erfahren Sie, wie Sie ein Amazon SageMaker AI Autopilot-Experiment erstellen, um verschiedene Modellkandidaten anhand eines tabellarischen Datensatzes zu untersuchen, vorzuverarbeiten und zu trainieren.

[Erkunden Sie den Leitfaden“](#)

- Automatisches Erstellen von Modellen für maschinelles Lernen

Erfahren Sie, wie Sie Amazon SageMaker AI Autopilot verwenden, um automatisch ein ML-Modell zu erstellen, zu trainieren und zu optimieren und das Modell einzusetzen, um Vorhersagen zu treffen.

[Beginnen Sie mit dem Tutorial“](#)

- Erkunden Sie die Modellierung mit Amazon SageMaker AI Autopilot anhand dieser Beispiel-Notebooks

Entdecken Sie Beispiel-Notebooks für Direktmarketing, zur Vorhersage von Kundenabwanderungen und wie Sie Ihren eigenen Datenverarbeitungscode in Amazon SageMaker AI Autopilot integrieren können.

[Erkunden Sie den Leitfaden“](#)

SageMaker AI Canvas

- Erste Schritte mit Amazon SageMaker AI Canvas

Erfahren Sie, wie Sie mit der Verwendung von SageMaker AI Canvas beginnen können.

[Erkunden Sie den Leitfaden“](#)

- Generieren Sie Vorhersagen für maschinelles Lernen, ohne Code zu schreiben

In diesem Tutorial wird erklärt, wie Sie Amazon SageMaker AI Canvas verwenden, um ML-Modelle zu erstellen und genaue Vorhersagen zu generieren, ohne eine einzige Codezeile schreiben zu müssen.

[Beginnen Sie mit dem Tutorial“](#)

- Tauchen Sie tiefer in SageMaker AI Canvas ein

Erfahren Sie mehr über SageMaker AI Canvas und seine visuellen ML-Funktionen ohne Code.

[Lesen Sie den Blog“](#)

- Verwenden Sie Amazon SageMaker AI Canvas, um Ihr erstes ML-Modell zu erstellen

Erfahren Sie, wie Sie mithilfe von Amazon SageMaker AI Canvas ein ML-Modell zur Bewertung der Kundenbindung auf der Grundlage einer E-Mail-Kampagne für neue Produkte und Dienstleistungen erstellen können.

[Fangen Sie mit dem Labor an“](#)

SageMaker AI Data Wrangler

- Erste Schritte mit Amazon SageMaker AI Data Wrangler

Erfahren Sie, wie SageMaker AI Data Wrangler eingerichtet wird, und finden Sie anschließend eine exemplarische Vorgehensweise anhand eines vorhandenen Beispieldatensatzes.

[Erkunden Sie den Leitfaden“](#)

- Bereiten Sie Trainingsdaten mit minimalem Code für maschinelles Lernen vor

Erfahren Sie, wie Sie mit Amazon SageMaker AI Data Wrangler Daten für ML vorbereiten.

[Beginnen Sie mit dem Tutorial“](#)

- SageMaker Tiefer Tauchworkshop von AI Data Wrangler

Erfahren Sie, wie Sie geeignete Analysetypen auf Ihren Datensatz anwenden, um Anomalien und Probleme zu erkennen, die abgeleiteten Ergebnisse results/insights zur Formulierung von Abhilfemaßnahmen im Zuge von Transformationen an Ihrem Datensatz zu verwenden und die richtige Auswahl und Reihenfolge der Transformationen mithilfe der von AI Data Wrangler bereitgestellten Schnellmodellierungsoptionen zu testen. SageMaker

[Fangen Sie mit dem Workshop an“](#)

SageMaker AI Ground Truth

- Erste Schritte mit Amazon Ground Truth

Erfahren Sie, wie Sie mit der Konsole einen Etikettierungsauftrag erstellen, eine öffentliche oder private Belegschaft zuweisen und den Etikettierungsauftrag an Ihre Belegschaft senden können. Erfahren Sie, wie Sie den Fortschritt eines Etikettierungsauftrags überwachen können.

[Erkunden Sie den Leitfaden“](#)

- Trainingsdaten für Machine Learning kennzeichnen

Erfahren Sie, wie Sie in Amazon SageMaker AI Ground Truth einen Labeling-Job einrichten, um Trainingsdaten für Ihr ML-Modell zu kommentieren.

[Beginnen Sie mit dem Tutorial“](#)

- Erste Schritte mit Amazon Ground Truth Plus Erfahren Sie, wie Sie die notwendigen Schritte ausführen, um ein Amazon SageMaker AI Ground Truth Plus-Projekt zu starten, Labels zu überprüfen und die Voraussetzungen für SageMaker AI Ground Truth Plus zu erfüllen.

[Erkunden Sie den Leitfaden“](#)

- Erste Schritte mit Amazon Ground Truth Sehen Sie sich an, wie Sie mit der Kennzeichnung Ihrer Daten in wenigen Minuten über die SageMaker KI Ground Truth Konsole beginnen können.

[Sehen Sie sich das Video an“](#)

- Amazon SageMaker AI Ground Truth Plus — Trainingsdatensätze ohne Code oder interne Ressourcen erstellen

Erfahren Sie mehr über Ground Truth Plus, einen schlüsselfertigen Service, der mithilfe von Fachkräften schnell hochwertige Trainingsdatensätze bereitstellt und die Kosten um bis zu 40 Prozent senkt.

[Lesen Sie den Blog“](#)

SageMaker AI JumpStart

- Beginnen Sie mit maschinellem Lernen mit SageMaker KI JumpStart

Entdecken Sie SageMaker JumpStart KI-Lösungsvorlagen, mit denen eine Infrastruktur für gängige Anwendungsfälle eingerichtet wird, und ausführbare Beispiel-Notebooks für maschinelles Lernen mit SageMaker KI.

[Erkunden Sie den Leitfaden“](#)

- Mit Amazon SageMaker AI können Sie schnell mit Ihrem Machine-Learning-Projekt beginnen JumpStart

Erfahren Sie, wie Sie Ihr ML-Projekt mithilfe von vortrainierten Modellen und vorgefertigten Lösungen von Amazon AI beschleunigen können. SageMaker JumpStart Anschließend können Sie das ausgewählte Modell über Amazon SageMaker AI Studio-Notebooks bereitstellen.

[Beginnen Sie mit dem Tutorial“](#)

- Machen Sie mit diesem Immersion Day-Workshop praktische Erfahrungen JumpStart mit Amazon SageMaker AI

Erfahren Sie, wie die Low-Code-ML-Funktionen von Amazon SageMaker AI Data Wrangler, Autopilot und Jumpstart es einfacher machen, schneller zu experimentieren und hochgenaue Modelle in die Produktion zu bringen.

[Fangen Sie mit dem Workshop an“](#)

SageMaker AI Pipelines

- Erste Schritte mit Amazon SageMaker AI Pipelines

Erfahren Sie, wie Sie end-to-end Workflows zur Verwaltung und Bereitstellung von SageMaker KI-Jobs erstellen. SageMaker AI Pipelines wird mit einer SageMaker AI Python SDK-Integration geliefert, sodass Sie jeden Schritt Ihrer Pipeline mithilfe einer Python-basierten Schnittstelle erstellen können.

[Erkunden Sie den Leitfaden“](#)

- Automatisieren Sie Arbeitsabläufe für maschinelles Lernen

Erfahren Sie, wie Sie mithilfe von Amazon SageMaker AI Pipelines, Amazon AI Model Registry und Amazon AI Clarify Workflows für end-to-end maschinelles Lernen (ML) erstellen und automatisieren. SageMaker SageMaker

[Beginnen Sie mit dem Tutorial“](#)

- So erstellen Sie vollautomatische ML-Workflows mit Amazon SageMaker AI Pipelines

Erfahren Sie mehr über Amazon SageMaker AI Pipelines, den weltweit ersten CI/CD ML-Service, der für jeden Entwickler und Datenwissenschaftler zugänglich ist. SageMaker AI Pipelines bringt CI/CD Pipelines zu ML und reduziert so die benötigte Codierungszeit.

[Sehen Sie sich das Video an](#)

SageMaker AI Studio

- Erstellen und trainieren Sie lokal ein Modell für maschinelles Lernen

Erfahren Sie, wie Sie ein ML-Modell lokal in Ihrem Amazon SageMaker AI Studio-Notizbuch erstellen und trainieren.

[Beginnen Sie mit dem Tutorial](#)

- SageMaker AI Studio-Integration mit EMR-Workshop

Erfahren Sie, wie Sie verteilte Verarbeitung in großem Maßstab nutzen können, um Daten aufzubereiten und anschließend Modelle für maschinelles Lernen zu trainieren.

[Beginnen Sie mit dem Workshop](#)

Die dritte Gruppe von AWS-Service KI/ML-Pfaden bietet ausführliche Dokumentation, praktische Tutorials und Ressourcen für die ersten Schritte mit AWS Trainium AWS Inferentia, und Amazon Titan.

AWS Trainium

- Skalierung verteilter Schulungen mit AWS Trainium Amazon EKS

Erfahren Sie, wie Sie von der allgemeinen Verfügbarkeit von Amazon EC2 Trn1-Instances profitieren können, die auf einem speziell entwickelten ML-Beschleuniger basieren, der optimiert wurde, um eine leistungsstarke, kostengünstige und massiv skalierbare Plattform für das Training von AWS Trainium Deep-Learning-Modellen in der Cloud bereitzustellen.

[Lesen Sie den Blog](#)

- Überblick über AWS Trainium

Erfahren Sie mehr über AWS Trainium den Beschleuniger für maschinelles Lernen (ML) der zweiten Generation, der AWS speziell für das Deep-Learning-Training von Modellen mit mehr als 100 B Parametern entwickelt wurde. Jede Amazon Elastic Compute Cloud (EC2) Trn1-Instance stellt bis zu 16 AWS Trainium Beschleuniger bereit, um eine leistungsstarke, kostengünstige Lösung für Deep-Learning-Schulungen (DL) in der Cloud bereitzustellen.

[Erkunden Sie den Leitfaden“](#)

- Empfohlene Trainium-Instances

Erfahren Sie, wie AWS Trainium Instances so konzipiert sind, dass sie hohe Leistung und Kosteneffizienz für Workloads mit Deep-Learning-Modellinferenz bieten.

[Erkunden Sie den Leitfaden“](#)

AWS Inferentia

- Überblick über AWS Inferentia

Erfahren Sie, wie Beschleuniger so konzipiert sind, dass sie hohe Leistung AWS zu den niedrigsten Kosten für Ihre Deep-Learning-Inferenzanwendungen (DL) bieten.

[Erkunden Sie den Leitfaden“](#)

- AWS Inferentia 2 baut auf AWS Inferentia 1 auf und bietet einen viermal höheren Durchsatz und eine zehnmal geringere Latenz

Erfahren Sie, wofür AWS Inferentia 2 optimiert ist — und erfahren Sie, wie es von Grund auf so konzipiert wurde, dass es eine höhere Leistung bietet LLMs und gleichzeitig die Kosten für generative KI-Inferenz senkt.

[Lesen Sie den Blog“](#)

- Inferenz für maschinelles Lernen mit AWS Inferentia

Erfahren Sie, wie Sie einen Amazon EKS-Cluster mit Knoten erstellen, auf denen Amazon EC2 Inf1-Instances ausgeführt werden, und (optional) eine Beispielanwendung bereitstellen. Amazon

EC2 Inf1-Instances werden mit AWS Inferentia Chips betrieben, die speziell für hohe Leistung und kostengünstigste Inferenz in der Cloud entwickelt wurden. AWS

[Erkunden Sie den Leitfaden“](#)

Amazon Titan

- Überblick über Amazon Titan

Erfahren Sie, wie Amazon Titan FMs mit großen Datensätzen vorab trainiert wurde, sodass es zu leistungsstarken Allzweckmodellen wird. Erfahren Sie, wie Sie sie unverändert — oder privat — verwenden können, um sie mit Ihren eigenen Daten für eine bestimmte Aufgabe anzupassen, ohne große Datenmengen mit Anmerkungen versehen zu müssen.

[Erkunden Sie den Leitfaden“](#)

Erkunden

- Architekturdiagramme

Diese Referenzarchitekturdiagramme zeigen Beispiele für verwendete AWS KI- und ML-Dienste.

[Erkunden Sie Architekturdiagramme“](#)

- Whitepapers

Schauen Sie sich Whitepapers an, die Ihnen den Einstieg erleichtern, und lernen Sie bewährte Verfahren bei der Auswahl und Nutzung von Diensten kennen. AI/ML

[Entdecken Sie Whitepapers“](#)

- AWS Lösungen

Entdecken Sie geprüfte Lösungen und Architekturanleitungen für gängige Anwendungsfälle für KI- und ML-Services.

[Entdecken Sie Lösungen“](#)

Ressourcen

Modelle der Stiftung

Zu den unterstützten Stiftungsmodellen gehören:

- [Der anthropische Claude](#)
- [Zusammenführen von Befehlen und Einbetten](#)
- [AI21 Labore Jurassic](#)
- [Meta Lama](#)
- [Mistral KI](#)
- [Stabile Diffusion XL](#)
- [Amazon Titan](#)

Mit Amazon Bedrock können Sie mit einer Vielzahl von Foundation-Modellen experimentieren und diese privat mit Ihren Daten anpassen.

Anwendungsfall oder branchenspezifische Services

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

Dazugehörige Blogbeiträge

- [Signifikante neue Funktionen erleichtern die Verwendung von Amazon Bedrock, um generative KI-Anwendungen zu erstellen und zu skalieren — und beeindruckende Ergebnisse zu erzielen](#)
- [AWS Inferentia und AWS Trainium Bereitstellung der niedrigsten Kosten für die Bereitstellung von Lama-3-Modellen in Amazon AI SageMaker JumpStart](#)

- [Revolutionieren Sie die Kundenzufriedenheit mit maßgeschneiderten Prämienmodellen für Ihr Unternehmen auf Amazon AI SageMaker](#)
- [Amazon Personalize bringt neue Rezepte auf den Markt, die größere Artikelkataloge mit geringerer Latenz unterstützen](#)

Dokumentverlauf

In der folgenden Tabelle werden die wichtigen Änderungen an diesem Entscheidungsleitfaden beschrieben. Für Benachrichtigungen über Aktualisierungen dieses Handbuchs können Sie einen RSS-Feed abonnieren.

Änderung	Beschreibung	Datum
Kleines Update	Aktualisierte Inhalte für Amazon Q und den neuesten KI- und ML-Stack von Amazon.	3. Mai 2024
Erstversion	Erste Veröffentlichung des Entscheidungsleitfadens.	24. Juli 2023

Die vorliegende Übersetzung wurde maschinell erstellt. Im Falle eines Konflikts oder eines Widerspruchs zwischen dieser übersetzten Fassung und der englischen Fassung (einschließlich infolge von Verzögerungen bei der Übersetzung) ist die englische Fassung maßgeblich.