Guida all'esame (AIF-C01)

AWS Certified AI Practitioner

AWS Certified AI Practitioner: Guida all'esame (AIF-C01)

Table of Contents

AWS Certified AI Practitioner (AIF-C01)

L'esame AWS Certified AI Practitioner (AIF-C01) è rivolto a persone che desiderano dimostrare una comprensione di base dei concetti dell'IA e degli strumenti di IA di AWS. Questa certificazione è incentrata sulle applicazioni pratiche dell'IA in ambito aziendale.

Argomenti

Introduzione

L'esame AWS Certified AI Practitioner (AIF-C01) è rivolto a persone che desiderano dimostrare una comprensione di base dei concetti dell'IA e degli strumenti di IA di AWS. Questa certificazione è incentrata sulle applicazioni pratiche dell'IA in ambito aziendale.

Inoltre, durante l'esame viene valutata la capacità dei candidati di completare le seguenti attività:

- Descrivere concetti, strategie e metodi relativi a ML, IA e IA generativa (GenAI), in generale e in AWS.
- Identificare l'uso appropriato delle tecnologie di IA/ML e GenAI per risolvere i problemi aziendali.
- Determinare i tipi corretti di tecnologie di IA/ML da applicare a specifici casi d'uso.

- Utilizzare le tecnologie di IA, ML e GenAI in modo responsabile.

Descrizione del candidato target

Il candidato target deve avere fino a 6 mesi di esperienza con le tecnologie di IA/ML in AWS. Il candidato target utilizza, ma non necessariamente crea, soluzioni di IA/ML in AWS.

Conoscenze su AWS consigliate

Il candidato target deve possedere le seguenti conoscenze su AWS:

- Familiarità con i principali servizi AWS (ad esempio, Amazon EC2, Amazon S3, AWS Lambda, Amazon Bedrock e Amazon SageMaker AI) e con i casi d'uso di tali servizi AWS
- Familiarità con il modello di responsabilità condivisa di AWS per la sicurezza e la conformità nel cloud AWS
- Familiarità con AWS Identity and Access Management (IAM) per proteggere e controllare l'accesso alle risorse AWS
- Familiarità con i modelli di prezzo dei servizi AWS

Attività lavorative che non rientrano nelle competenze del candidato target

Il seguente elenco illustra le attività lavorative che non rientrano nelle competenze previste per il candidato target. L'elenco non è esaustivo. Le seguenti attività non rientrano nell'ambito dell'esame:

- Sviluppo o codifica di algoritmi o modelli di IA/ML
- Implementazione di tecniche di ingegneria dei dati o ingegneria delle caratteristiche
- Esecuzione del tuning degli iperparametri o dell'ottimizzazione dei modelli
- Creazione e distribuzione di infrastrutture o pipeline di IA/ML
- Esecuzione di analisi matematiche o statistiche dei modelli di IA/ML
- Implementazione di protocolli di sicurezza o conformità per i sistemi di IA/ML
- Sviluppo e implementazione di policy e framework di governance per le soluzioni di IA/ML

Contenuto dell'esame

Tipi di domande

L'esame include uno o più dei seguenti tipi di domande:

- Scelta multipla: una risposta corretta e tre risposte errate (distrattori).
- Risposta multipla: due o più risposte corrette su cinque o più opzioni di risposta. È necessario selezionare tutte le risposte corrette per ricevere i crediti per la domanda.
- Ordinamento: un elenco di 3-5 risposte per completare un'attività specifica. È necessario selezionare le risposte corrette e posizionarle nell'ordine giusto per ricevere i crediti per la domanda.
- Abbinamento: un elenco di risposte da abbinare a un elenco di 3-7 prompt. È necessario abbinare correttamente tutte le coppie per ricevere i crediti per la domanda.

Le domande senza risposta vengono valutate come errate. Non è prevista alcuna penalità per chi prova a indovinare. L'esame prevede 50 domande che influiscono sul punteggio finale.

Contenuto senza punteggio

L'esame include 15 domande alle quali non viene assegnato un punteggio e che non influiscono sul risultato finale. AWS raccoglie informazioni sulle prestazioni relativamente a queste domande, al fine di valutare la possibilità di convertirle in futuro in domande a punteggio. In sede di esame, le domande che non influiscono sul punteggio non verranno distinte dalle altre.

Risultati dell'esame

L'esame AWS Certified AI Practitioner (AIF-C01) prevede un esito netto, superamento o mancato superamento. La valutazione avviene in base a uno standard minimo stabilito da professionisti AWS che seguono le best practice e le linee guida del settore delle certificazioni.

I risultati dell'esame sono espressi da un punteggio compreso tra 100 e 1.000. Il punteggio minimo richiesto per il superamento della prova è 700. Il punteggio riflette le prestazioni complessive del candidato all'esame e indica se l'esame è stato superato o meno. I modelli di punteggio dimensionato aiutano a equiparare i punteggi tra moduli dell'esame, i quali potrebbero presentare livelli di difficoltà leggermente diversi.

Il report relativo al punteggio può contenere una tabella di classificazione del rendimento in ogni sezione. Per l'esame viene impiegato un modello di punteggio compensativo; ciò significa che non è necessario ottenere un punteggio sufficiente in ogni sezione. L'esame viene superato se il punteggio complessivo ottenuto corrisponde almeno al minimo richiesto.

Ogni sezione ha un proprio peso specifico, quindi alcune di esse presentano più domande di altre. La tabella delle classificazioni include informazioni generali che evidenziano i punti forti e deboli del candidato. Interpreta con la massima attenzione il feedback relativo a ogni sezione.

Descrizione del contenuto

Questa guida all'esame include informazioni sui pesi, sui domini del contenuto, sulle attività e sulle competenze per l'esame. Non fornisce un elenco esaustivo dei contenuti dell'esame.

Di seguito sono elencati i domini del contenuto e i pesi dell'esame:

- Dominio del contenuto 1: Fondamenti di IA e ML (20% dei contenuti a punteggio)
- Dominio del contenuto 2: Fondamenti di GenAI (24% dei contenuti a punteggio)
- Dominio del contenuto 3: Applicazioni dei modelli di fondazione (28% dei contenuti a punteggio)
- Dominio del contenuto 4: Linee guida per un'IA responsabile (14% dei contenuti a punteggio)
- Dominio del contenuto 5: Sicurezza, conformità e governance per le soluzioni di IA (14% dei contenuti a punteggio)

Dominio del contenuto 1: Fondamenti di IA e ML

Il dominio 1 tratta i principi fondamentali dell'intelligenza artificiale (IA) e del machine learning (ML) e rappresenta il 20% dei contenuti a punteggio dell'esame.

Attività

- Obiettivo 1.1: Spiegazione dei termini e dei concetti di base correlati all'IA
- Obiettivo 1.2: Identificazione dei casi d'uso pratici per l'IA
- Obiettivo 1.3: Descrizione del ciclo di vita dello sviluppo delle tecnologie di IA/ ML

Obiettivo 1.1: Spiegazione dei termini e dei concetti di base correlati all'IA

Obiettivi:

- Definire i termini di base correlati all'IA, ad esempio IA, ML, deep learning, reti neurali, visione artificiale, elaborazione del linguaggio naturale, modello, algoritmo, addestramento e inferenza, bias, equità, adattamento, modello linguistico di grandi dimensioni (LLM), IA generativa (GenAI), IA agentica.
- Descrivere somiglianze e differenze tra IA, ML, GenAI, deep learning e IA agentica.
- Descrivere vari tipi di inferenza, ad esempio in batch, in tempo reale, asincrona, serverless.
- Descrivere i diversi tipi di dati nei modelli di IA (ad esempio, etichettati e non etichettati, tabellari, serie temporali, immagini, testo, strutturati e non strutturati).
- Descrivere diversi tipi di apprendimento IA/ML, ad esempio metodi di apprendimento supervisionato, apprendimento senza supervisione, apprendimento per rinforzo.

Obiettivo 1.2: Identificazione dei casi d'uso pratici per l'IA

Obiettivi:

- Riconoscere le applicazioni in cui le soluzioni di IA/ML possono generare valore (ad esempio, assistenza ai processi decisionali umani, scalabilità delle soluzioni e automazione).
- Determinare quando le soluzioni di IA/ML non sono appropriate (ad esempio, analisi costi-benefici o situazioni in cui è necessario ottenere un risultato specifico anziché una previsione).
- Selezionare le tecniche di IA/ML appropriate per casi d'uso specifici, ad esempio regressione, classificazione, clustering.
- Individuare esempi di applicazioni reali di IA, ad esempio visione artificiale, elaborazione del linguaggio naturale, riconoscimento vocale, sistemi di suggerimento, rilevamento frodi, esecuzione di previsioni, basi di conoscenza, IA agentica.
- Spiegare le funzionalità dei servizi di IA/ML gestiti da AWS (ad esempio, Amazon SageMaker AI, Amazon Transcribe, Amazon Translate, Amazon Comprehend, Amazon Lex e Amazon Polly).
- Individuare quando sia appropriato applicare modelli ML tradizionali rispetto a modelli di fondazione (FM) a un caso d’uso specifico, ad esempio sulla base di considerazioni normative, requisiti di spiegabilità o vincoli operativi.

Obiettivo 1.3: Descrizione del ciclo di vita dello sviluppo delle tecnologie di IA/ ML

Obiettivi:

- Descrivere e differenziare i componenti di una pipeline di IA/ML.
- Descrivere le origini dei modelli FM, ad esempio modelli pre-addestrati open source, modelli personalizzati di addestramento.
- Descrivere i metodi per utilizzare un modello in produzione (ad esempio, servizio API gestito o API self-hosted).
- Individuare servizi e funzionalità AWS pertinenti per ogni fase di una pipeline di IA/ML, ad esempio Amazon Bedrock, Amazon Q, Amazon Quick, Kiro,SageMaker AI.
- Descrivere i concetti fondamentali delle operazioni di ML (MLOps), ad esempio, sperimentazione, processi ripetibili, sistemi scalabili, gestione del debito tecnico, raggiungimento del livello di preparazione necessario per la produzione, monitoraggio dei modelli e riaddestramento dei modelli.
- Descrivere le metriche di prestazione dei modelli, ad esempio accuratezza, precisione, richiamo e punteggio F1, nonché le metriche aziendali, ad esempio costo per utente, costi di sviluppo, feedback dei clienti, ritorno sull'investimento (ROI), per valutare i modelli di ML.

Dominio del contenuto 2: Fondamenti di GenAI

Il dominio 2 tratta i principi fondamentali dell'intelligenza artificiale generativa (GenAI) e rappresenta il 24% dei contenuti a punteggio dell'esame.

Attività

- Obiettivo 2.1: Spiegazione dei concetti di base correlati all’IA generativa
- Obiettivo 2.2: Comprensione delle funzionalità e dei limiti delle applicazioni di GenAI per risolvere i problemi aziendali
- Obiettivo 2.3: Descrizione dell'infrastruttura e delle tecnologie AWS per la creazione di applicazioni di GenAI

Obiettivo 2.1: Spiegazione dei concetti di base correlati all’IA generativa

Obiettivi:

- Definire i concetti fondamentali correlati alla GenAI, ad esempio token, chunking, embedding, vettori, progettazione dei prompt, modelli linguistici di grandi dimensioni (LLM) basati su trasformatore, modelli di fondazione (FM), modelli multimodali, modelli di diffusione.

- Identificare i potenziali casi d'uso per i modelli di GenAI, ad esempio, generazione di immagini, video e audio, sintesi, assistenti di IA, traduzione, generazione di codice, agenti per il servizio clienti, ricerche e motori di suggerimenti.
- Descrivere il ciclo di vita degli FM, ad esempio selezione dei dati, selezione dei modelli, pre-addestramento, fine-tuning, valutazione, distribuzione, feedback.
- Descrivere il modello di prezzo basato su token e il suo effetto su costi e prestazioni per l'inferenza.
- Descrivere il ruolo dell'ingegneria del contesto nelle applicazioni che utilizzano FM.
- Definire i concetti fondamentali correlati all'IA agentica, ad esempio modelli di sistema multiagente per applicazioni di IA complesse, Model Context Protocol (MCP) e il suo ruolo nel collegamento degli agenti a sistemi esterni, modelli di comunicazione multiagente, gestione della memoria, utilizzo degli strumenti e orchestrazione dei flussi di lavoro.

Obiettivo 2.2: Comprensione delle funzionalità e dei limiti delle applicazioni di GenAI per risolvere i problemi aziendali

Obiettivi:

- Descrivere i vantaggi derivanti dall'IA generativa, ad esempio adattabilità, reattività, capacità conversazionale, capacità di generare contenuti.
- Identificare gli svantaggi delle soluzioni di IA generativa, ad esempio allucinazioni, interpretabilità, imprecisione e non determinismo.
- Individuare i fattori da tenere in considerazione per la scelta di modelli di IA generativa, ad esempio tipi di modello, requisiti prestazionali, capacità, vincoli, conformità, costi, latenza, complessità dei modelli.
- Stabilire le metriche e il valore di business delle applicazioni di IA generativa, ad esempio prestazioni multidominio, ROI, efficienza, tasso di conversione, ricavi medi per utente, accuratezza, valore medio del cliente.

Obiettivo 2.3: Descrizione dell'infrastruttura e delle tecnologie AWS per la creazione di applicazioni di GenAI

Obiettivi:

- Individuare servizi e funzionalità AWS per sviluppare applicazioni di IA generativa, ad esempio Amazon Bedrock, Amazon SageMaker AI, SageMaker JumpStart, Amazon Quick, Kiro,Strands Agents, Amazon Bedrock AgentCore.
- Descrivere i vantaggi dell'utilizzo dei servizi di IA generativa AWS per creare applicazioni, ad esempio accessibilità, barriere all'ingresso ridotte, efficienza, convenienza, time-to-market ridotto e capacità di raggiungere gli obiettivi aziendali.
- Descrivere i vantaggi dell'infrastruttura AWS per le applicazioni di IA generativa, ad esempio, sicurezza, conformità e responsabilità.
- Descrivere i compromessi in termini di costo relativi ai servizi di IA generativa AWS, ad esempio reattività, disponibilità, ridondanza, prestazioni, copertura regionale, determinazione dei prezzi basata su token, throughput di provisioning e modelli personalizzati.

Dominio del contenuto 3: Applicazioni dei modelli di fondazione

Il dominio 3 tratta le applicazioni dei modelli di fondazione e rappresenta il 28% dei contenuti a punteggio dell'esame.

Attività

- Obiettivo 3.1: Illustrazione delle considerazioni di progettazione per le applicazioni che utilizzano modelli di fondazione (FM)
- Obiettivo 3.2: Scelta delle tecniche efficaci per la progettazione dei prompt
- Obiettivo 3.3: Descrizione del processo di addestramento e fine-tuning per gli FM
- Obiettivo 3.4: Descrizione dei metodi per valutare le prestazioni degli FM

Obiettivo 3.1: Illustrazione delle considerazioni di progettazione per le applicazioni che utilizzano modelli di fondazione (FM)

Obiettivi:

- Individuare i criteri per la scelta di FM, ad esempio costi, modalità, latenza, modelli multilingue, dimensioni del modello, complessità del modello, personalizzazione, lunghezza di input/output, caching dei prompt.
- Descrivere l'effetto dei parametri di inferenza sulle risposte dei modelli, ad esempio, temperatura e lunghezza di input/output.

- Definire la generazione potenziata da recupero dati (RAG) e descrivere le relative applicazioni aziendali, ad esempio Knowledge Base per Amazon Bedrock.
- Identificare i servizi AWS che consentono di memorizzare gli embedding all'interno di database vettoriali, ad esempio Servizio OpenSearch di Amazon, Amazon Aurora, Amazon Neptune e Amazon RDS per PostgreSQL.
- Spiegare i compromessi in termini di costo relativi ai vari approcci alla personalizzazione degli FM, ad esempio pre-addestramento, fine-tuning, apprendimento contestuale, RAG e distillazione di modelli.
- Definire il ruolo degli agenti IA e descriverne le applicazioni aziendali.

Obiettivo 3.2: Scelta delle tecniche efficaci per la progettazione dei prompt

Obiettivi:

- Definire i concetti e i costrutti della progettazione dei prompt (ad esempio, contesto, istruzioni, prompt negativi).
- Definire le tecniche per la progettazione dei prompt (ad esempio, catena di pensiero, zero-shot, single-shot, few-shot e template di prompt).
- Identificare e descrivere i vantaggi e le best practice per la progettazione dei prompt (ad esempio, miglioramento della qualità della risposta, sperimentazione, guardrail, individuazione, specificità e concisione e utilizzo di più commenti).
- Definire i potenziali rischi e limiti della progettazione dei prompt (ad esempio, esposizione, poisoning, hijacking e jailbreaking).
- Descrivere il versionamento dei prompt e le strategie di gestione che utilizzano Amazon Bedrock Prompt Management.

Obiettivo 3.3: Descrizione del processo di addestramento e fine-tuning per gli FM

Obiettivi:

- Descrivere gli elementi chiave dell'addestramento di un FM (ad esempio, pre-addestramento, fine-tuning, pre-addestramento continuo e distillazione).
- Definire i metodi per il fine-tuning di un FM (ad esempio, tuning delle istruzioni, adattamento dei modelli a domini specifici, apprendimento per trasferimento e pre-addestramento continuo).

- Descrivere come preparare i dati per il fine-tuning di un FM, ad esempio data curation, governance, dimensioni, etichettatura, rappresentatività, apprendimento per rinforzo da feedback umano (RLHF).

Obiettivo 3.4: Descrizione dei metodi per valutare le prestazioni degli FM

Obiettivi:

- Stabilire approcci per valutare le prestazioni degli FM, ad esempio valutazione human-in-the-loop (HITL), set di dati di benchmark, valutazione del modello in Amazon Bedrock.
- Individuare le metriche pertinenti per valutare le prestazioni degli FM, ad esempio Recall-Oriented Understudy for Gisting Evaluation (ROUGE), valutazione bilingue (BLEU), BERTScore, LLM-as-a-judge.
- Stabilire se un FM soddisfa in modo efficace gli obiettivi aziendali, ad esempio produttività, coinvolgimento degli utenti, progettazione delle attività.
- Individuare approcci per valutare le prestazioni delle applicazioni create con FM, ad esempio RAG, agenti, flussi di lavoro.
- Individuare metriche di allineamento degli obiettivi aziendali per applicazioni di IA, ad esempio percentuale di completamento delle attività, soddisfazione degli utenti, costo per interazione.

Dominio del contenuto 4: Linee guida per un'IA responsabile

Il dominio 4 tratta le linee guida per un'IA responsabile e rappresenta il 14% dei contenuti a punteggio dell'esame.

Attività

- Obiettivo 4.1: Spiegazione dello sviluppo di sistemi di IA responsabili
- Obiettivo 4.2: Riconoscimento dell'importanza di modelli trasparenti e spiegabili

Obiettivo 4.1: Spiegazione dello sviluppo di sistemi di IA responsabili

Obiettivi:

- Identificare le caratteristiche dell'IA responsabile (ad esempio, bias, equità, inclusività, robustezza, sicurezza e veridicità).

- Spiegare come utilizzare gli strumenti per identificare le caratteristiche dell'IA responsabile (ad esempio, Guardrail per Amazon Bedrock).
- Definire pratiche responsabili per selezionare un modello (ad esempio, considerazioni ambientali e sostenibilità).
- Individuare i rischi legali correlati all'utilizzo di soluzioni di IA generativa, ad esempio reclami per violazioni della proprietà intellettuale, output di modelli con bias, perdita di fiducia dei clienti, rischi per l'utente finale, allucinazioni.
- Identificare le caratteristiche dei set di dati (ad esempio, inclusività, varietà, origini dati curate e set di dati bilanciati).
- Descrivere gli effetti dei bias e della varianza (ad esempio, effetti sui gruppi demografici, imprecisione, overfitting e underfitting).
- Descrivere gli strumenti per rilevare e monitorare i bias, l'affidabilità e la veridicità, ad esempio analisi della qualità delle etichette, verifiche condotte da umani, analisi dei sottogruppi, Amazon SageMaker Clarify, SageMaker Model Monitor e Amazon Augmented AI (Amazon A2I).

Obiettivo 4.2: Riconoscimento dell'importanza di modelli trasparenti e spiegabili

Obiettivi:

- Descrivere le differenze tra i modelli che sono trasparenti e spiegabili e quelli che non lo sono.
- Descrivere gli strumenti per individuare modelli trasparenti e spiegabili, ad esempio Amazon SageMaker Model Cards, SageMaker Clarify, valutazioni dei modelli in Amazon Bedrock, modelli open source, dati, licenze.
- Identificare i compromessi tra sicurezza e trasparenza dei modelli (ad esempio, misurazione di interpretabilità e prestazioni).
- Descrivere i principi della progettazione umanocentrica per un'IA spiegabile, ad esempio meccanismi di feedback degli utenti, trasparenza delle decisioni dell'IA.

Dominio del contenuto 5: Sicurezza, conformità e governance per le soluzioni di IA

Il dominio 5 tratta la sicurezza, la conformità e la governance per le soluzioni di IA e rappresenta il 14% dei contenuti a punteggio dell'esame.

Attività

- Obiettivo 5.1: Spiegazione dei metodi per proteggere i sistemi di IA
- Obiettivo 5.2: Riconoscimento delle normative in materia di governance e conformità per i sistemi di IA

Obiettivo 5.1: Spiegazione dei metodi per proteggere i sistemi di IA

Obiettivi:

- Individuare servizi e funzionalità AWS per proteggere i sistemi di IA, ad esempio ruoli, policy e autorizzazioni IAM; crittografia; Amazon Macie; AWS PrivateLink; modello di responsabilità condivisa di AWS; Amazon Bedrock AgentCore Identity; Policy in AgentCore; Guardrail per Amazon Bedrock.
- Descrivere il concetto di identificazione delle origini e documentazione delle origini dei dati (ad esempio, data lineage, catalogazione dei dati e Amazon SageMaker Model Cards).
- Descrivere le best practice per un'ingegneria dei dati sicura (ad esempio, valutazione della qualità dei dati, implementazione di tecnologie volte al miglioramento della privacy, controllo dell'accesso ai dati e integrità dei dati).
- Descrivere le considerazioni in materia di sicurezza e privacy per i sistemi di IA, ad esempio sicurezza delle applicazioni, rilevamento delle minacce, gestione delle vulnerabilità, protezione dell'infrastruttura, iniezione di prompt, crittografia dei dati inattivi e in transito, prevenzione di data leakage, filtraggio e convalida degli output, requisiti dei percorsi di verifica e registrazione dei log per le interazioni con l'IA, tossicità.
- Descrivere i metodi di rilevamento delle allucinazioni e le tecniche di grounding per migliorare l'accuratezza degli output, ad esempio grounding con generazione potenziata da recupero dati (RAG), convalida degli output, determinazione del punteggio di confidenza.

Obiettivo 5.2: Riconoscimento delle normative in materia di governance e conformità per i sistemi di IA

Obiettivi:

- Identificare i servizi e le funzionalità AWS per agevolare la governance e la conformità normativa (ad esempio, AWS Config, Amazon Inspector, Gestione audit AWS, AWS Artifact, AWS CloudTrail e AWS Trusted Advisor).

- Descrivere le strategie di governance dei dati (ad esempio, cicli di vita dei dati, registrazione di log, residenza, monitoraggio, osservazione e conservazione).
- Descrivere i processi per seguire i protocolli di governance (ad esempio, policy, cadenza delle revisioni, strategie di revisione, framework di governance come la Generative AI Security Scoping Matrix, standard di trasparenza e requisiti di formazione dei team).

Servizi AWS trattati in sede di esame

Il seguente elenco contiene i servizi e le funzionalità AWS trattati nell'esame AWS Certified AI Practitioner (AIF-C01). L'elenco non è esaustivo ed è soggetto a modifiche. Le offerte AWS sono suddivise in categorie in base alle loro funzioni principali.

Argomenti

Analisi

- Scambio dati su AWS
- Amazon EMR
- AWS Glue
- AWS Glue DataBrew
- AWS Lake Formation

- Servizio OpenSearch di Amazon
- Amazon Quick
- Amazon Redshift

Gestione finanziaria del cloud

- AWS Budgets
- AWS Cost Explorer

Calcolo

- Amazon EC2
- AWS Lambda

Container

- Amazon Elastic Container Service (Amazon ECS)
- Amazon Elastic Kubernetes Service (Amazon EKS)

Database

- Amazon Aurora
- Amazon DocumentDB (compatibile con MongoDB)
- Amazon DynamoDB
- Amazon ElastiCache
- Amazon Neptune
- Amazon RDS

Strumenti di sviluppo

- Kiro
- Strands Agents

- Amazon Q

Machine learning

- Amazon Augmented AI (Amazon A2I)
- Amazon Bedrock
- Amazon Bedrock AgentCore
- Amazon Comprehend
- Amazon Kendra
- Amazon Lex
- Amazon Nova
- Amazon Personalize
- Amazon Polly
- Amazon Rekognition
- Amazon SageMaker AI
- Amazon SageMaker JumpStart
- Amazon Textract
- Amazon Transcribe
- Amazon Translate
- AWS Transform

Gestione e governance su AWS

- AWS CloudTrail
- Amazon CloudWatch
- AWS Config
- AWS Trusted Advisor
- Strumento AWS Well-Architected

Reti e distribuzione di contenuti

- Amazon CloudFront

- Amazon VPC

Sicurezza, identità e conformità

- AWS Artifact
- Gestione audit AWS
- AWS Identity and Access Management (IAM)
- Amazon Inspector
- Servizio AWS di gestione delle chiavi (AWS KMS)
- Amazon Macie
- AWS Secrets Manager

Archiviazione

- Amazon S3
- Amazon S3 Glacier

Servizi AWS non trattati in sede di esame

Il seguente elenco contiene i servizi e le funzionalità AWS non trattati nell'esame. L'elenco non è esaustivo ed è soggetto a modifiche. Dall'elenco sono escluse le offerte AWS del tutto estranee ai ruoli target per l'esame:

Argomenti

- Analisi
- Integrazione di applicazioni
- Applicazioni aziendali
- Gestione finanziaria del cloud
- Calcolo
- Container
- Abilitazione dei clienti
- Database

- Strumenti di sviluppo
- Servizi informatici per utenti finali
- Frontend per il web e i dispositivi mobili
- Internet of Things
- Machine learning
- Gestione e governance su AWS
- Servizi multimediali
- Migrazione e trasferimento
- Reti e distribuzione di contenuti
- Sicurezza, identità e conformità
- Archiviazione

Analisi

- AWS Clean Rooms
- Amazon CloudSearch
- Streaming gestito da Amazon per Apache Kafka (Amazon MSK)

Integrazione di applicazioni

- Amazon AppFlow
- Amazon MQ
- Amazon Simple Workflow Service (Amazon SWF)

Applicazioni aziendali

- Amazon Chime
- Amazon Pinpoint
- Amazon Simple Email Service (Amazon SES)
- AWS Supply Chain
- AWS Wickr
- Amazon WorkMail

Gestione finanziaria del cloud

- AWS Application Cost Profiler
- AWS Billing Conductor
- Marketplace AWS

Calcolo

- AWS App Runner
- AWS Elastic Beanstalk
- EC2 Image Builder
- Amazon Lightsail

Container

- Servizio Red Hat OpenShift su AWS (ROSA)

Abilitazione dei clienti

- AWS IQ
- Servizi gestiti AWS (AMS)
- AWS re:Post Private
- Supporto AWS

Database

- Amazon Keyspaces (per Apache Cassandra)
- Database Amazon Quantum Ledger (Amazon QLDB)
- Amazon Timestream

Strumenti di sviluppo

- AWS AppConfig

- Strumento AWS per la creazione di applicazioni
- AWS CloudShell
- Amazon CodeCatalyst
- AWS CodeStar
- AWS Fault Injection Service
- AWS X-Ray

Servizi informatici per utenti finali

- Amazon AppStream 2.0
- Amazon WorkSpaces
- Amazon WorkSpaces Thin Client
- Amazon WorkSpaces Web

Frontend per il web e i dispositivi mobili

- AWS Amplify
- AWS AppSync
- AWS Device Farm
- Servizio di posizione Amazon

Internet of Things

- AWS IoT Analytics
- AWS IoT Core
- AWS IoT Device Defender
- Gestione del dispositivo AWS IoT
- AWS IoT Events
- AWS IoT FleetWise
- FreeRTOS
- AWS IoT GreenGrass
- AWS IoT 1-Click

- AWS IoT RoboRunner
- AWS IoT SiteWise
- AWS IoT TwinMaker

Machine learning

- AWS DeepComposer
- AWS HealthImaging
- AWS HealthOmics
- Amazon Monitron
- AWS Panorama

Gestione e governance su AWS

- AWS Control Tower
- Dashboard AWS Health
- Avvio della procedura guidata AWS
- Strumento AWS di gestione delle licenze
- Grafana gestito da Amazon
- Servizio gestito da Amazon per Prometheus
- AWS OpsWorks
- AWS Organizations
- AWS Proton
- AWS Resilience Hub
- Esploratore di risorse AWS
- Gruppi di risorse AWS
- Strumento di gestione degli incidenti AWS Systems Manager
- Catalogo dei servizi AWS
- Service Quotas
- Gestione reti di telecomunicazioni di AWS
- Notifiche AWS agli utenti

Servizi multimediali

- Amazon Elastic Transcoder
- AWS Elemental MediaConnect
- AWS Elemental MediaConvert
- AWS Elemental MediaLive
- AWS Elemental MediaPackage
- AWS Elemental MediaStore
- AWS Elemental MediaTailor
- Servizio video interattivo Amazon (Amazon IVS)
- Amazon Nimble Studio

Migrazione e trasferimento

- Servizio AWS di individuazione delle applicazioni
- Servizio AWS di migrazione delle applicazioni
- AWS Database Migration Service (AWS DMS)
- AWS DataSync
- Modernizzazione del mainframe AWS
- Hub di migrazione AWS
- Famiglia AWS Snow
- AWS Transfer Family

Reti e distribuzione di contenuti

- AWS App Mesh
- AWS Cloud Map
- AWS Direct Connect
- AWS Global Accelerator
- 5G privato di AWS
- Amazon Route 53

- Amazon Route 53 Application Recovery Controller
- Amazon VPC IP Address Manager (IPAM)

Sicurezza, identità e conformità

- Gestione certificati AWS (ACM)
- AWS CloudHSM
- Amazon Cognito
- Amazon Detective
- Servizio di directory AWS
- Gestione dei firewall AWS
- Amazon GuardDuty
- Centro identità AWS IAM
- AWS Payment Cryptography
- Autorità di certificazione privata AWS
- AWS Resource Access Manager (AWS RAM)
- Centrale di sicurezza AWS
- Amazon Security Lake
- AWS Shield
- AWS Signer
- Autorizzazioni verificate da Amazon
- AWS WAF

Archiviazione

- Backup AWS
- AWS Elastic Disaster Recovery

Revisioni

Le guide agli esami AWS vengono riviste e aggiornate periodicamente al fine di garantire che gli esami di certificazione abbiano a oggetto le competenze, i servizi e le funzionalità AWS pertinenti

per i ruoli lavorativi a cui è destinata una certificazione. I contenuti aggiornati vengono introdotti negli esami circa un mese dopo la pubblicazione degli aggiornamenti alle guide.

Argomenti

Cronologia delle modifiche

Versione	Data di pubblicazione
1.0	26 marzo 2026
1.1	30 aprile 2026

Modifiche agli obiettivi

Versione 1.0	Versione 1.1
Obiettivo 1.1.1: Definire i termini di base correlati all'IA, ad esempio IA, ML, deep learning, reti neurali, visione artificiale, elaborazione del linguaggio naturale, modello, algoritmo, addestramento e inferenza, bias, equità, adattamento e modelli linguistici di grandi dimensioni (LLM).	Obiettivo 1.1.1: Definire i termini di base correlati all'IA, ad esempio IA, ML, deep learning, reti neurali, visione artificiale, elaborazione del linguaggio naturale, modello, algoritmo, addestramento e inferenza, bias, equità, adattamento, modello linguistico di grandi dimensioni (LLM), IA generativa (GenAI), IA agentica.
Obiettivo 1.1.2: Descrivere somiglianze e differenze tra IA, ML, GenAI e deep learning.	Obiettivo 1.1.2: Descrivere somiglianze e differenze tra IA, ML, GenAI, deep learning e IA agentica.

Versione 1.0	Versione 1.1
Obiettivo 1.1.3: Descrivere vari tipi di inferenza (ad esempio, batch e in tempo reale).	Obiettivo 1.1.3: Descrivere vari tipi di inferenza, ad esempio in batch, in tempo reale, asincrona, serverless.
Obiettivo 1.1.5: Descrivere l'apprendimento supervisionato, l'apprendimento senza supervisione e l'apprendimento per rinforzo.	Obiettivo 1.1.5: Descrivere diversi tipi di apprendimento IA/ML, ad esempio metodi di apprendimento supervisionato, apprendim ento senza supervisione, apprendimento per rinforzo.
Obiettivo 1.2.4: Identificare esempi di applicazi oni reali di IA (ad esempio, visione artificiale, elaborazione del linguaggio naturale, riconosci mento vocale, sistemi di suggerimento, rilevamento frodi ed esecuzione di previsioni).	Obiettivo 1.2.4: Individuare esempi di applicazi oni reali di IA, ad esempio visione artificia le, elaborazione del linguaggio naturale, riconoscimento vocale, sistemi di suggerime nto, rilevamento frodi, esecuzione di previsioni, basi di conoscenza, IA agentica.
Obiettivo 1.3.1: Descrivere i componenti di una pipeline di ML (ad esempio, raccolta dei dati, analisi esplorativa dei dati, pre-elaborazione dei dati, ingegneria delle caratteristiche, addestram ento dei modelli, tuning degli iperparametri, valutazione, distribuzione e monitoraggio).	Obiettivo 1.3.1: Descrivere e differenziare i componenti di una pipeline di IA/ML.
Obiettivo 1.3.4: Identificare i servizi e le funzionalità AWS pertinenti per ogni fase di una pipeline di ML (ad esempio, SageMaker AI, SageMaker Data Wrangler, SageMaker Feature Store e SageMaker Model Monitor).	Obiettivo 1.3.4: Individuare servizi e funzionalità AWS pertinenti per ogni fase di una pipeline di IA/ML, ad esempio Amazon Bedrock, Amazon Q, Amazon Quick, Kiro,SageMaker AI.

Versione 1.0	Versione 1.1
Obiettivo 1.3.6: Descrivere le metriche di prestazione dei modelli, ad esempio accuratezza, area sotto la curva (AUC) e punteggio F1, e le metriche aziendali, ad esempio costo per utente, costi di sviluppo, feedback dei clienti, ritorno sull'investimento (ROI), per valutare i modelli di ML.	Obiettivo 1.3.6: Descrivere le metriche di prestazione dei modelli, ad esempio accuratezza, precisione, richiamo e punteggio F1, nonché le metriche aziendali, ad esempio costo per utente, costi di sviluppo, feedback dei clienti, ritorno sull'investimento (ROI), per valutare i modelli di ML.
Obiettivo 2.2.1: Descrivere i vantaggi delle soluzioni di GenAI (ad esempio, adattabilità, reattività e semplicità).	Obiettivo 2.2.1: Descrivere i vantaggi derivanti dall'IA generativa, ad esempio adattabilità, reattività, capacità conversazionale, capacità di generare contenuti.
Obiettivo 2.2.3: Identificare i fattori da considerare per la selezione dei modelli di GenAI (ad esempio, tipi di modello, requisiti prestazionali, capacità, vincoli e conformità).	Obiettivo 2.2.3: Individuare i fattori da tenere in considerazione per la scelta di modelli di IA generativa, ad esempio tipi di modello, requisiti prestazionali, capacità, vincoli, conformità, costi, latenza, complessità dei modelli.
Obiettivo 2.2.4: Determinare le metriche e il valore di business per le applicazioni di GenAI (ad esempio, prestazioni multidominio, efficienza, tasso di conversione, ricavi medi per utente, accuratezza e valore medio del cliente).	Obiettivo 2.2.4: Stabilire le metriche e il valore di business delle applicazioni di IA generativa, ad esempio prestazioni multidominio, ROI, efficienza, tasso di conversione, ricavi medi per utente, accuratezza, valore medio del cliente.
Obiettivo 2.3.1: Identificare i servizi e le funzionalità AWS per sviluppare applicazioni di GenAI (ad esempio, Amazon SageMaker JumpStart, Amazon Bedrock PartyRock, Amazon Q e Amazon Bedrock Data Automation).	Obiettivo 2.3.1: Individuare servizi e funzionalità AWS per sviluppare applicazioni di IA generativa, ad esempio Amazon Bedrock, Amazon SageMaker AI, SageMaker JumpStart, Amazon Quick, Kiro,Strands Agents, Amazon Bedrock AgentCore.

Versione 1.0	Versione 1.1
Obiettivo 3.1.5: Spiegare i compromessi di costo dei vari approcci alla personalizzazione degli FM (ad esempio, pre-addestramento, fine-tuning, apprendimento contestuale e RAG).	Obiettivo 3.1.5: Spiegare i compromessi in termini di costo relativi ai vari approcci alla personalizzazione degli FM, ad esempio pre-addestramento, fine-tuning, apprendimento contestuale, RAG e distillazione di modelli.
Obiettivo 3.1.6: Descrivere il ruolo degli agenti nelle attività in più fasi (ad esempio, Agent per Amazon Bedrock, IA agentica e Model Context Protocol).	Obiettivo 3.1.6: Definire il ruolo degli agenti IA e descriverne le applicazioni aziendali.
Obiettivo 3.4.1: Determinare gli approcci per valutare le prestazioni degli FM (ad esempio, valutazione umana, set di dati di benchmark e valutazione del modello in Amazon Bedrock).	Obiettivo 3.4.1: Stabilire approcci per valutare le prestazioni degli FM, ad esempio valutazio ne human-in-the-loop (HITL), set di dati di benchmark, valutazione del modello in Amazon Bedrock.
Obiettivo 3.4.2: Identificare le metriche pertinenti per valutare le prestazioni degli FM, ad esempio Recall-Oriented Understudy for Gisting Evaluation (ROUGE) valutazione bilingue (BLEU) e BERTScore.	Obiettivo 3.4.2: Individuare le metriche pertinenti per valutare le prestazioni degli FM, ad esempio Recall-Oriented Understudy for Gisting Evaluation (ROUGE), valutazione bilingue (BLEU), BERTScore, LLM-as-a-judge.
Obiettivo 4.2.2: Descrivere gli strumenti per identificare modelli trasparenti e spiegabili (ad esempio, Amazon SageMaker Model Cards, modelli open source, dati, licenze).	Obiettivo 4.2.2: Descrivere gli strumenti per individuare modelli trasparenti e spiegabili, ad esempio Amazon SageMaker Model Cards, SageMaker Clarify, valutazioni dei modelli in Amazon Bedrock, modelli open source, dati, licenze.
Obiettivo 4.2.4: Descrivere i principi della progettazione umanocentrica per un'IA spiegabile.	Obiettivo 4.2.4: Descrivere i principi della progettazione umanocentrica per un'IA spiegabile, ad esempio meccanismi di feedback degli utenti, trasparenza delle decisioni dell'IA.

Versione 1.0	Versione 1.1
Obiettivo 5.1.1: Identificare i servizi e le funzionalità AWS per proteggere i sistemi di IA (ad esempio, ruoli, policy e autorizzazioni IAM, crittografia, Amazon Macie, AWS PrivateLink, modello di responsabilità condivisa di AWS).	Obiettivo 5.1.1: Individuare servizi e funzionalità AWS per proteggere i sistemi di IA, ad esempio ruoli, policy e autorizzazioni IAM; crittografia; Amazon Macie; AWS PrivateLink; modello di responsabilità condivisa di AWS; Amazon Bedrock AgentCore Identity; Policy in AgentCore; Guardrail per Amazon Bedrock.
Obiettivo 5.1.4: Illustrare le considerazioni in materia di sicurezza e privacy per i sistemi di IA (ad esempio, sicurezza delle applicazioni, rilevamento delle minacce, gestione delle vulnerabilità, protezione dell'infrastruttura, iniezione di prompt, crittografia a riposo e in transito).	Obiettivo 5.1.4: Descrivere le considerazioni in materia di sicurezza e privacy per i sistemi di IA, ad esempio sicurezza delle applicazioni, rilevamento delle minacce, gestione delle vulnerabilità, protezione dell'infrastruttura, iniezione di prompt, crittografia dei dati inattivi e in transito, prevenzione di data leakage, filtraggio e convalida degli output, requisiti dei percorsi di verifica e registrazione dei log per le interazioni con l'IA, tossicità.

Obiettivi aggiunti

- Obiettivo 1.2.6: Individuare quando sia appropriato applicare modelli ML tradizionali rispetto a modelli di fondazione (FM) a un caso d’uso specifico, ad esempio sulla base di considerazioni normative, requisiti di spiegabilità, vincoli operativi.
- Obiettivo 2.1.4: Descrivere il modello di prezzo basato su token e il suo effetto su costi e prestazioni per l'inferenza.
- Obiettivo 2.1.5: Descrivere il ruolo dell'ingegneria del contesto nelle applicazioni che utilizzano FM.
- Obiettivo 2.1.6: Definire i concetti fondamentali correlati all'IA agentica, ad esempio modelli di sistema multiagente per applicazioni di IA complesse, Model Context Protocol (MCP) e il suo ruolo nel collegamento degli agenti a sistemi esterni, modelli di comunicazione multiagente, gestione della memoria, utilizzo degli strumenti e orchestrazione dei flussi di lavoro.
- Obiettivo 3.2.5: Descrivere il versionamento dei prompt e le strategie di gestione che utilizzano Amazon Bedrock Prompt Management.

- Obiettivo 3.4.5: Individuare metriche di allineamento degli obiettivi aziendali per applicazioni di IA, ad esempio percentuale di completamento delle attività, soddisfazione degli utenti, costo per interazione.
- Obiettivo 5.1.5: Descrivere i metodi di rilevamento delle allucinazioni e le tecniche di grounding per migliorare l'accuratezza degli output, ad esempio grounding con generazione potenziata da recupero dati (RAG), convalida degli output, determinazione del punteggio di confidenza.

Modifiche ai servizi trattati e non trattati in sede d'esame

Servizi aggiunti all'elenco degli argomenti trattati in sede d'esame

- Amazon Aurora
- Amazon Bedrock AgentCore
- Kiro
- Strands Agents
- Amazon Q
- Amazon SageMaker JumpStart
- AWS Transform

Servizi rimossi dall'elenco degli argomenti trattati in sede d'esame

- Amazon MemoryDB

Servizi rimossi dall'elenco degli argomenti non trattati in sede d'esame

- AWS DeepComposer
- Amazon FinSpace
- Amazon Honeycode
- Centro identità AWS IAM
- Marketplace AWS
- AWS Organizations
- Amazon WorkDocs

Sondaggio

Quanto è stata utile questa guida all'esame? Per farci sapere cosa ne pensi, partecipa al sondaggio.