

Développer des applications d'IA génératives sur AWS

Date et durée
Code formation : AWS07FR Durée : 2 jours Nombre d'heures : 14 heures
Description
Pour innover en IA générative, AWS propose Amazon Bedrock, qui facilite l'accès aux modèles de fondation, et Amazon CodeWhisperer, un assistant de codage précieux. Ces services permettent aux entreprises de créer des applications intelligentes et d'accroître leur productivité. En complément, AWS propose une gamme étendue de services d'IA pour enrichir les capacités de l'IA générative, couvrant des domaines comme l'analyse d'images ou la transcription vocale. Cette formation intensive de 2 jours est conçue pour vous doter des compétences pratiques indispensables à la maîtrise de l'IA générative et des services AWS qui l'accompagnent. Vous plongerez au cœur des cas d'utilisation concrets, explorerez en profondeur les principes de l'ingénierie des prompts et aborderez les considérations cruciales liées à l'IA responsable. Grâce à une série d'ateliers pratiques interactifs, vous apprendrez à concevoir des prompts efficaces, à évaluer avec pertinence les résultats obtenus et à planifier stratégiquement l'implémentation de vos futurs projets d'IA générative. À l'issue de ce programme complet, vous aurez développé l'expertise nécessaire pour comprendre, appliquer et exploiter l'IA générative de manière responsable et efficace sur l'écosystème AWS. Vous maîtriserez les concepts clés, les meilleures pratiques et les outils essentiels, tels qu'Amazon Bedrock et LangChain, pour tirer pleinement parti du potentiel de cette technologie transformative au sein de votre organisation.
Objectifs
À l'issue de cette formation en IA générative sur AWS, vous atteindrez les objectifs de compétences suivants : • maîtriser les concepts fondamentaux de l'IA générative, y compris son importance, ses bénéfices, ses risques et sa terminologie clé ; • planifier un projet d'IA générative en identifiant la valeur métier des cas d'usage, en expliquant les étapes et en gérant les risques ; • comprendre le fonctionnement d'Amazon Bedrock, ses avantages, ses cas d'usage typiques, son architecture et sa structure de coûts, et réaliser une démonstration pratique ; • appliquer les principes et techniques de l'ingénierie des prompts, y compris les méthodes de base et avancées, la sélection de modèles spécifiques et l'atténuation des biais ; • développer des applications d'IA générative en identifiant leurs composants, en personnalisant les modèles de fondation (FMs) avec leurs APIs et paramètres d'inférence ; • Intégrer LangChain avec de grands modèles de langage (LLMs) pour construire des applications complexes, y compris l'utilisation de chaînes, de modèles de chat, d'embeddings, de chargeurs de documents, de récupérateurs et d'agents ; • mettre en œuvre des architectures d'IA générative et appliquer les concepts pour construire et tester des cas d'usage avec Amazon Bedrock, LangChain et l'approche RAG (Retrieval Augmented Generation) ; • identifier les services AWS pour le monitoring, la sécurisation et la gouvernance des applications Amazon Bedrock.

Points forts
• Formateurs experts AWS certifiés : bénéficiez de l'expertise de formateurs reconnus et certifiés par AWS, qui possèdent une connaissance approfondie des pratiques d'IA générative et des outils et services AWS pour le développement et le déploiement d'applications basées sur les grands modèles de langage. • Mise en pratique interactive : maîtrisez les outils et techniques de l'IA générative sur AWS grâce à des démonstrations concrètes et des ateliers pratiques. Vous serez ainsi préparé à relever les défis réels de l'ingénierie des prompts, de l'intégration avec LangChain, de l'utilisation d'Amazon Bedrock, et de la mise en œuvre d'architectures d'applications d'IA générative dans le cloud. • Acquisition de compétences clés : le contenu de la formation est soigneusement conçu pour vous permettre d'acquérir les compétences essentielles en matière de conception et d'implémentation de solutions d'IA générative, de gestion des grands modèles de langage et de LangChain, et de gestion efficace des ressources sur AWS.
Modalités d'évaluation
Quiz / QCM Travaux Pratiques
Pré-requis
<i>Suivre cette formation AWS nécessite les prérequis recommandés suivants :</i> • avoir suivi la formation AWS Technical Essentials ; • posséder une maîtrise de niveau intermédiaire du langage Python.
Public
<i>Cette formation s'adresse aux publics suivants :</i> • les développeurs de logiciels intéressés par l'exploitation de grands modèles de langage sans avoir besoin de les ajuster (fine-tuning).
Programme
Module 1 : comprendre l'IA générative (l'art du possible) • Qu'est-ce que le machine learning (ML) ? • Les bases de l'IA générative et ses concepts fondamentaux. • Les cas d'usage de l'IA générative. • Les risques et les bénéfices de l'IA. • L'IA générative en pratique : aperçu général. Module 2 : planifier et contextualiser un projet d'IA • Le contexte de l'IA générative pour l'entreprise. • Les étapes clés de planification d'un projet d'intégration de l'IA. • L'identification et l'atténuation des risques spécifiques aux projets d'IA générative. Module 3 : démarrer avec Amazon Bedrock • Présentation du service Amazon Bedrock.

- L'architecture et les cas d'usage.
- Les bonnes pratiques pour utiliser Amazon Bedrock.
- La configuration des accès et l'utilisation des Playgrounds (*démo*).

Module 4 : maîtriser les bases de l'ingénierie des prompts

- Les bases des modèles de fondation.
- Les fondamentaux de l'ingénierie des prompts.
- Les techniques de prompt de base et avancées.
- L'affinement d'un prompt texte de base (*démo*).
- Les techniques de prompt spécifiques aux modèles.
- La gestion des utilisations abusives des prompts.
- L'atténuation des biais.
- L'atténuation des biais d'image (*démo*).

Module 5 : construire les composants d'application Amazon Bedrock

- Les applications et les cas d'usage.
- Les composants principaux d'application.
- Les modèles de fondation et l'interface FM.
- L'utilisation des jeux de données et des embeddings.
- Les composants d'application supplémentaires.
- La génération augmentée par récupération (RAG).
- Le fine-tuning de modèle.
- La sécurisation des applications d'IA générative.
- L'architecture des applications d'IA générative.
- L'intégration de mots (*démo*).

Module 6 : utiliser les modèles de fondation Amazon Bedrock

- Présentation des modèles de fondation.
- L'utilisation des FMs Amazon Bedrock pour l'inférence.
- Les méthodes Amazon Bedrock.
- La protection des données et l'auditabilité.

Lab :

- Invoquer le modèle Amazon Bedrock pour la génération de texte en utilisant un prompt "zero-shot".

Module 7 : intégrer le Framework LangChain

- L'optimisation des performances des LLM.
- L'intégration d'AWS et du LangChain.
- L'utilisation de modèles avec LangChain.
- La construction des prompts.
- La structuration des documents avec des indexes.
- Le stockage et la récupération des données avec la mémoire.
- L'utilisation des chaînes pour séquencer les composants.
- La gestion des ressources externes avec les agents LangChain.

Module 8 : appliquer les modèles d'architecture

- Les bases des modèles d'architecture.
- Le résumé de texte.
- La question-réponse.
- Les chatbots.

- La génération de code.
- Le LangChain et les agents pour Amazon Bedrock.

Labs :

- Utiliser Amazon Titan Text Premier pour résumer le texte de petits fichiers.
- Résumer de longs textes avec Amazon Titan.
- Utiliser Amazon Bedrock pour la question-réponse.
- Construire un chatbot.
- Utiliser les modèles Amazon Bedrock pour la génération de code.
- Construire des applications conversationnelles avec l'API Converse.

AWS, Amazon Bedrock, Amazon Titan et les autres marques AWS sont des marques déposées d'Amazon.com, inc. ou de ses filiales.