

Développement d'Applications d'IA Générative sur AWS

Durée: 2 jours

Public cible: Professionnel TI

À qui s'adresse cette formation ?

Personnes développant des logiciels intéressées par l'utilisation des LLMs sans ajustement fin

Objectifs de la formation

- Décrire l'IA générative et son alignement avec l'apprentissage automatique
- Définir l'importance de l'IA générative et expliquer ses risques et avantages potentiels
- Identifier la valeur commerciale des cas d'utilisation de l'IA générative
- Discuter des bases techniques et de la terminologie clé de l'IA générative
- Expliquer les étapes de planification d'un projet d'IA générative
- Identifier certains des risques et des mesures d'atténuation lors de l'utilisation de l'IA générative
- Comprendre le fonctionnement d'Amazon Bedrock
- Se familiariser avec les concepts de base d'Amazon Bedrock
- Reconnaître les avantages d'Amazon Bedrock
- Lister les cas d'utilisation typiques d'Amazon Bedrock
- Décrire l'architecture typique associée à une solution Amazon Bedrock
- Comprendre la structure des coûts d'Amazon Bedrock
- Mettre en œuvre une démonstration d'Amazon Bedrock dans la console de gestion AWS
- Définir l'ingénierie des prompts et appliquer les meilleures pratiques générales lors de l'interaction avec les modèles de base (FMs)
- Identifier les types de techniques de prompt de base, y compris l'apprentissage zéro-shot et few-shot
- Appliquer des techniques de prompt avancées lorsque nécessaire pour votre cas d'utilisation
- Identifier les techniques de prompt les mieux adaptées à des modèles spécifiques
- Identifier les utilisations abusives potentielles des prompts
-

Analyser les biais potentiels dans les réponses des FMs et concevoir des prompts qui atténuent ces biais

- Identifier les composants d'une application d'IA générative et comment personnaliser un FM
- Décrire les modèles de base d'Amazon Bedrock, les paramètres d'inférence et les principales API d'Amazon Bedrock
- Identifier les offres d'Amazon Web Services (AWS) qui aident à surveiller, sécuriser et gouverner vos applications Amazon Bedrock
- Décrire comment intégrer LangChain avec les LLMs, les modèles de prompt, les chaînes, les modèles de chat, les modèles d'embeddings de texte, les chargeurs de documents, les récupérateurs et les agents pour Amazon Bedrock
- Décrire les modèles d'architecture que vous pouvez implémenter avec Amazon Bedrock pour construire des applications d'IA générative
- Appliquer les concepts pour construire et tester des cas d'utilisation exemplaires utilisant les différents modèles d'Amazon Bedrock, LangChain, et l'approche de génération augmentée par récupération (RAG)

Description sommaire

Ce cours est conçu pour introduire l'intelligence artificielle générative (IA) aux personnes développant des logiciels intéressées par l'utilisation de grands modèles de langage (LLMs) sans ajustement fin. Le cours offre une vue d'ensemble de l'IA générative, la planification d'un projet d'IA générative, les bases d'Amazon Bedrock, les fondamentaux de l'ingénierie des prompts, et les modèles d'architecture pour construire des applications d'IA générative en utilisant Amazon Bedrock et LangChain.

Contenu du plan de cours

Module 1 : Introduction à l'IA Générative - L'Art du Possible

- Vue d'ensemble de l'apprentissage automatique
- Bases de l'IA générative
- Cas d'utilisation de l'IA générative
-

L'IA générative en pratique

- Risques et avantages

Module 2 : Planification d'un Projet d'IA Générative

- Fondamentaux de l'IA générative
- L'IA générative en pratique
- Contexte de l'IA générative
- Étapes de planification d'un projet d'IA générative
- Risques et mesures d'atténuation

Module 3 : Démarrage avec Amazon Bedrock

- Introduction à Amazon Bedrock
- Architecture et cas d'utilisation
- Comment utiliser Amazon Bedrock
- Démonstration : Configuration de l'accès à Bedrock et utilisation des playgrounds
- Bases des modèles de base
- Fondamentaux de l'ingénierie des prompts
- Techniques de prompt de base
- Techniques de prompt avancées
- Techniques de prompt spécifiques aux modèles
- Démonstration : Ajustement fin d'un prompt textuel de base
- Aborder les utilisations abusives des prompts
- Atténuer les biais
- Démonstration : Atténuation des biais d'image

Module 4 : Fondamentaux de l'Ingénierie des Prompts

- Bases des modèles de base
- Fondamentaux de l'ingénierie des prompts
-

Techniques de prompt de base

- Techniques de prompt avancées
- Techniques de prompt spécifiques aux modèles
- Démonstration : Ajustement fin d'un prompt textuel de base
- Aborder les utilisations abusives des prompts
- Atténuer les biais • Démonstration : Atténuation des biais d'image

Module 5 : Composants des Applications Amazon Bedrock

- • Vue d'ensemble des composants des applications d'IA générative
- • Modèles de base et interface FM
- • Travailler avec des ensembles de données et des embeddings
- • Démonstration : Embeddings de mots
- • Composants supplémentaires des applications
- • Génération augmentée par récupération (RAG)
- • Ajustement fin des modèles
- • Sécuriser les applications d'IA générative
- • Architecture des applications d'IA générative

Module 6 : Modèles de Base Amazon Bedrock

- Introduction aux modèles de base Amazon Bedrock
- Utilisation des FMs d'Amazon Bedrock pour l'inférence
- Méthodes d'Amazon Bedrock
- Protection des données et auditabilité
- Démonstration : Invocation du modèle Bedrock pour la génération de texte en utilisant un prompt zéro-shot

Module 7 : LangChain

- Optimisation des performances des LLMs

- Utilisation des modèles avec LangChain
- Construction de prompts
- Démonstration : Bedrock avec LangChain en utilisant un prompt incluant du contexte
- Structuration des documents avec des index
- Stockage et récupération des données avec la mémoire
- Utilisation des chaînes pour séquencer les composants
- Gestion des ressources externes avec les agents LangChain

Module 8 : Modèles d'Architecture

- Introduction aux modèles d'architecture
- Résumé de texte
- Démonstration : Résumé de texte de petits fichiers avec Anthropic Claude
- Démonstration : Résumé de texte abstrait avec Amazon Titan utilisant LangChain
- Réponse aux questions
- Démonstration : Utilisation d'Amazon Bedrock pour répondre aux questions
- Chatbot • Démonstration : Interface conversationnelle - Chatbot avec LLM AI21
- Génération de code
- Démonstration : Utilisation des modèles Amazon Bedrock pour la génération de code
- LangChain et agents pour Amazon Bedrock
- Démonstration : Intégration des modèles Amazon Bedrock avec les agents LangChain

Approche et méthode pédagogique

Approche pratique et structurée combinant théorie ciblée et ateliers guidés. Les participants développent progressivement des applications d'IA générative sur AWS à travers des exercices concrets inspirés de scénarios professionnels, favorisant une mise en application immédiate des apprentissages. Ils apprennent à intégrer des services d'IA générative, concevoir des solutions intelligentes et exploiter les capacités du cloud AWS pour créer des applications innovantes. Animée par un for-

mateur certifié, la formation privilégie l'interactivité et le développement de compétences techniques directement transférables pour concevoir et déployer des solutions d'IA en contexte professionnel.

Prérequis

- Avoir complété AWS Technical Essentials ou contenu équivalent
- Maîtrise intermédiaire de Python

Recommandations

- Bases solides en programmation (Python recommandé)
- Compréhension des concepts d'intelligence artificielle et machine learning
- Familiarité avec les API, services cloud et intégration d'applications
- Connaissance des services AWS (EC2, S3, SageMaker, Lambda – atout important)
- Notions d'IA générative et de modèles de langage (prompts, embeddings – atout)