

Développement avancé d'IA générative sur AWS

Durée: 3 Jours Réf de cours: GK910033

Résumé:

Maîtriser la mise en œuvre de solutions d'IA générative prêtes à l'emploi sur AWS.

La formation « Développement avancé d'IA générative sur AWS » s'adresse aux développeurs souhaitant maîtriser la mise en œuvre de solutions d'IA générative prêtes à l'emploi sur AWS. Elle répond aux besoins des organisations qui se lancent dans l'aventure de l'IA générative et explique comment élaborer des stratégies complètes en matière d'IA générative, en adéquation avec les objectifs commerciaux généraux.

Cette formation avancée de trois jours, animée par un formateur, permet d'acquérir une expertise sur l'ensemble de la gamme des technologies d'IA générative, des modèles de base aux modèles d'intégration d'entreprise. Vous découvrirez les techniques avancées de traitement des données, la mise en œuvre de bases de données vectorielles et l'augmentation de la récupération, l'ingénierie et la gouvernance sophistiquées des prompts, les systèmes d'IA agentique et l'intégration d'outils, les mesures de sûreté et de sécurité de l'IA, les stratégies d'optimisation des performances et de gestion des coûts, les solutions complètes de surveillance et d'observabilité, ainsi que les cadres de test et de validation.

Mis à jour 23/06/2026

Public visé:

Cette formation s'adresse aux développeurs confirmés et aux ingénieurs IT se spécialisant en IA générative.

Objectifs pédagogiques:

- A l'issue de la formation, les participants seront capables de :
- Développer des solutions d'IA générative prêtes à la production à l'aide des services AWS, répondant aux exigences des entreprises en matière de sécurité, d'évolutivité et de fiabilité
- Évaluer et sélectionner les modèles de base adaptés à des cas d'utilisation métier spécifiques, notamment en comparant les performances et en mettant en œuvre des architectures de sélection dynamique des modèles
- Concevoir et mettre en œuvre des systèmes de modèles de base résilients intégrant des disjoncteurs, un déploiement inter-régions et des stratégies de dégradation en douceur
- Construire des pipelines complets de traitement des données pour des entrées multimodales, incluant des workflows de validation et des techniques d'optimisation
- Mettre en œuvre des solutions sophistiquées de bases de données vectorielles à l'aide d'Amazon Bedrock Knowledge Bases, d'OpenSearch et d'approches hybrides pour une augmentation efficace de la recherche
- Créer et gérer des frameworks avancés d'ingénierie des prompts, incluant le raisonnement par chaîne de pensée et des systèmes de gouvernance des prompts à l'échelle de l'entreprise
- Développer des agents IA autonomes à l'aide d'Amazon Bedrock Agents, en mettant en œuvre des modèles de raisonnement complexes et des capacités d'intégration d'outils
- Mettre en place des contrôles complets en matière de sécurité et de sûreté de l'IA, notamment le filtrage de contenu, la protection de la vie privée et des mécanismes de tests adversariaux
- Optimiser les performances et gérer les coûts grâce à des stratégies d'efficacité des jetons, des implémentations par lots et des systèmes de mise en cache intelligents
- Concevoir et mettre en œuvre des solutions complètes de surveillance et d'observabilité pour les applications basées sur des modèles de base
- Créer des cadres de test et de validation systématiques pour l'assurance qualité continue des applications d'IA
- Intégrer des solutions d'IA générative au sein d'environnements d'entreprise à l'aide de modèles architecturaux sécurisés, conformes et évolutifs

Pré-requis:

Pour ce cours d'un niveau avancé, les participants doivent avoir les prérequis suivants :

- Connaissances techniques de base sur AWS telles qu'obtenues par le cours GK4501 AWS : Notions techniques de base Amazon Web Services ou expérience professionnelle équivalente
- Connaissances de base en IA générative sur AWS telles qu'obtenues par le cours GK910027 Essentiels de l'IA Générative sur AWS ou expérience professionnelle équivalente
- Au moins 2 ans d'expérience dans le développement d'applications en production sur AWS ou avec des technologies open source, ainsi qu'une expérience générale en IA/ML ou en ingénierie des données
- 1 an d'expérience pratique dans la mise en œuvre de solutions d'IA générative

Test et certification

■

Après cette formation, nous vous conseillons le(s) module(s) suivant(s):

Pas de recommandation

Contenu:

Module 1: Sélection et configuration des modèles de base

- Cadrage d'évaluation des modèles de base d'entreprise
- Modèles d'architecture pour la sélection dynamique des modèles
- Conception de systèmes de modèles de base résilients
- Optimisation des coûts et modélisation économique

Module 2 : Traitement avancé des données pour les modèles de base

- Validation complète des données et assurance qualité
- Chaînes de traitement des données multimodales
- Optimisation des données d'entrée et amélioration des performances

Module 3 : Bases de données vectorielles et recherche augmentée

- Architecture des bases de données vectorielles d'entreprise
- Stratégies avancées de traitement et de découpage des documents
- Mise en œuvre de systèmes de recherche sophistiqués
- Travaux pratiques : Développer des applications de génération augmentée par la recherche (RAG) avec les bases de connaissances Amazon Bedrock

Module 4 : Ingénierie des prompts et gouvernance

- Cadres avancés d'ingénierie des prompts
- Systèmes complexes d'orchestration des prompts
- Gouvernance et gestion des prompts en entreprise
- Travaux pratiques : Développement d'un modèle de conversation avec les API Amazon Bedrock

Module 5 : IA agentique et intégration d'outils

- Architecture et évolution de l'IA agentique
- Mise en œuvre des agents Amazon Bedrock
- Écosystème de services AWS pour l'IA agentique
- Intégration d'outils et observabilité en production

Module 6 : Sécurité et sûreté de l'IA

- Mise en œuvre exhaustive de la sécurité des contenus
- Architecture d'IA préservant la confidentialité
- Cadres de gouvernance et de conformité de l'IA

Module 7 : Optimisation des performances et gestion des coûts

- Efficacité des jetons et optimisation des coûts
- Architecture système haute performance
- Mise en œuvre de systèmes de mise en cache intelligents
- Travaux pratiques : Création d'une IA générative sécurisée et responsable avec des garde-fous pour Amazon Bedrock

Module 8 : Surveillance et observabilité de l'IA générative

- Systèmes de surveillance des modèles de base
- Impact sur l'activité et gestion de la valeur
- Dépannage et diagnostics spécifiques à l'IA

Module 9 : Tests, validation et amélioration continue

- Cadres d'évaluation complets de l'IA
- Assurance qualité et amélioration continue
- Évaluation et optimisation des systèmes RAG

Module 10 : Modèles d'intégration d'entreprise

- Connectivité d'entreprise et architecture d'intégration
- Accès sécurisé et gestion des identités
- Déploiements inter-environnements et hybrides

Module 11 : Synthèse de la formation

- Résumé de la formation
- Prochaines étapes et ressources supplémentaires

Méthodes pédagogiques :

Un support de cours officiel sera fourni aux participants.

Autres moyens pédagogiques et de suivi:

- Compétence du formateur : Les experts qui animent la formation sont des spécialistes des matières abordées et ont au minimum cinq ans d'expérience d'animation. Nos équipes ont validé à la fois leurs connaissances techniques (certifications le cas échéant) ainsi que leur compétence pédagogique.
- Suivi d'exécution : Une feuille d'emargement par demi-journée de présence est signée par tous les participants et le formateur.
- En fin de formation, le participant est invité à s'auto-évaluer sur l'atteinte des objectifs énoncés, et à répondre à un questionnaire de satisfaction qui sera ensuite étudié par nos équipes pédagogiques en vue de maintenir et d'améliorer la qualité de nos prestations.

Délais d'inscription :