

Small Language Models (SLM) et Edge AI pour l'Entreprise — IA Locale, Rapide et Souveraine

Lien : https://maroc-performance.com/formation/small-language-models-slm-et-edge-ai-pour-lentreprise-ia-locale-rapide-et-souveraine

DURÉE
8 jours (56h)

RÉFÉRENCE
IEEI48

CATÉGORIE
IA pour Déploiement LLM, Architecture IA et MLOps

Maroc Performance

OBJECTIFS DE LA FORMATION

À l'issue de cette formation, vous serez capable de :

- Maîtriser les fondamentaux des Small Language Models et de l’Edge AI appliqués aux environnements d’entreprise.
- Comprendre les différences entre SLM, LLM Cloud et architectures hybrides afin de sélectionner la solution adaptée.
- Concevoir des architectures permettant d’exécuter des modèles IA localement avec une faible latence.
- Évaluer et sélectionner les SLM selon leurs performances, contraintes techniques et besoins métiers.
- Maîtriser les principales techniques de compression et d’optimisation des modèles IA pour l’Edge.
- Déployer des modèles capables de fonctionner localement avec une consommation maîtrisée des ressources.
- Connecter les SLM aux données internes de l’entreprise grâce au RAG et aux architectures de knowledge management.
- Concevoir des solutions d’IA souveraine protégeant les données sensibles et réduisant la dépendance aux infrastructures externes.
- Déployer des assistants et applications SLM adaptés aux processus métiers et industriels.
- Intégrer les SLM aux environnements IoT, IIoT, Edge et systèmes industriels.
- Mettre en place des dispositifs de monitoring, sécurité, gouvernance et gestion du cycle de vie des modèles.
- Construire une stratégie d’industrialisation SLM + Edge AI permettant de déployer une IA locale, rapide, souveraine et scalable.

Maroc Performance

👥 POUR QUI ?

- Chief Technology Officers (CTO), Chief AI Officers et responsables des stratégies technologiques et IA.
- Directeurs IT, responsables architecture et responsables transformation digitale.
- AI Engineers, Machine Learning Engineers et MLOps Engineers.
- Data Scientists et Data Engineers souhaitant développer et déployer des modèles IA localement.
- Cloud Engineers, Edge Engineers et architectes Cloud/Edge.
- Enterprise Architects, Solutions Architects et responsables des plateformes IA.
- Responsables cybersécurité, souveraineté numérique et gouvernance des données.
- Responsables IoT, IIoT, Industrie 4.0 et systèmes industriels connectés.
- Ingénieurs automatisme, OT, informatique industrielle et systèmes embarqués.
- Responsables innovation, R&D et développement de solutions d’intelligence artificielle.
- Consultants et experts en IA générative, Edge Computing, Machine Learning et transformation numérique.
- Dirigeants et managers souhaitant développer une capacité d’IA locale, souveraine et maîtrisée au sein de leur organisation.

Maroc Performance



Programme détaillé

1 / FONDAMENTAUX DES SMALL LANGUAGE MODELS ET DE L'EDGE AI

- Comprendre les principes des Small Language Models (SLM) et leurs différences avec les Large Language Models (LLM).
- Identifier les cas d'usage nécessitant une IA locale, rapide, spécialisée et faiblement consommatrice en ressources.
- Comprendre les avantages des architectures SLM + Edge AI en matière de latence, confidentialité, résilience et souveraineté.

2 / ARCHITECTURE SLM POUR L'IA LOCALE D'ENTREPRISE

- Concevoir une architecture intégrant modèles locaux, Edge Devices, serveurs sur site et systèmes d'information.
- Définir les interactions entre SLM, applications métiers, bases de données, API et infrastructures Cloud.
- Choisir une architecture adaptée aux contraintes de calcul, mémoire, connectivité et sécurité de l'entreprise.

3 / SÉLECTION ET ÉVALUATION DES SMALL LANGUAGE MODELS

- Comparer les SLM selon leur taille, performance, consommation de ressources et capacité de raisonnement.
- Évaluer les modèles selon les contraintes métier, linguistiques, techniques et réglementaires de l'organisation.
- Définir des critères de benchmark permettant de sélectionner le modèle le plus pertinent pour chaque cas d'usage.

4 / OPTIMISATION ET COMPRESSION DES MODÈLES IA

- Comprendre les techniques de quantization, pruning, distillation et compression des modèles.
- Réduire l'empreinte mémoire et la consommation de calcul tout en préservant la qualité des réponses.

Maroc Performance

- Optimiser les modèles pour leur exécution sur CPU, GPU, NPU et autres architectures Edge.

5 / INFÉRENCE LOCALE ET IA EN TEMPS RÉEL

- Déployer des SLM directement sur des infrastructures locales ou des équipements Edge.
- Réduire la latence et la dépendance à la connectivité Cloud grâce à l'inférence locale.
- Concevoir des applications capables de fonctionner efficacement dans des environnements à connectivité limitée ou intermittente.

6 / SLM ET DONNÉES D'ENTREPRISE

- Connecter les SLM aux bases de connaissances, documents et données internes de l'entreprise.
- Mettre en œuvre des architectures Retrieval-Augmented Generation (RAG) adaptées aux environnements locaux.
- Garantir la pertinence, la fraîcheur, la traçabilité et la confidentialité des informations utilisées par les modèles.

7 / IA SOUVERAINE ET PROTECTION DES DONNÉES

- Comprendre les enjeux de souveraineté numérique liés à l'hébergement et au traitement des données par l'IA.
- Concevoir des architectures permettant de conserver les données sensibles et les modèles dans l'environnement de l'entreprise.
- Mettre en place des mécanismes de contrôle des accès, chiffrement, journalisation et segmentation des environnements IA.

8 / SLM POUR LES PROCESSUS MÉTIERS ET L'ASSISTANCE INTELLIGENTE

- Déployer des assistants IA locaux pour les fonctions métiers, industrielles, commerciales, financières et administratives.
- Automatiser la classification, extraction, synthèse et génération de contenus à partir des données internes.
- Adapter les SLM aux besoins spécifiques de l'entreprise grâce au prompting, RAG et techniques d'adaptation des modèles.

9 / EDGE AI, IOT ET ENVIRONNEMENTS INDUSTRIELS

- Intégrer les SLM et l'Edge AI aux architectures IoT, IIoT, OT et systèmes industriels.

Maroc Performance

- Exploiter l’intelligence locale pour analyser les données des équipements et prendre des décisions à faible latence.
- Concevoir des applications combinant langage naturel, données capteurs, Computer Vision et automatisation industrielle.

10 / PERFORMANCE, MONITORING ET GESTION DU CYCLE DE VIE

- Mesurer la latence, le throughput, la consommation mémoire, l’utilisation CPU/GPU et la qualité des réponses.
- Mettre en place un monitoring permettant de détecter les dérives, erreurs et dégradations des modèles locaux.
- Organiser le versioning, la mise à jour, le rollback et la maintenance des SLM déployés sur les différents sites.

11 / SÉCURITÉ, GOUVERNANCE ET RISQUE DES SLM

- Identifier les risques liés aux hallucinations, aux attaques adversariales, aux données sensibles et aux modèles compromis.
- Mettre en place des mécanismes de validation, filtrage, contrôle des sorties et supervision humaine.
- Définir une gouvernance couvrant sécurité, conformité, traçabilité, responsabilité et gestion du cycle de vie des modèles.

12 / INDUSTRIALISATION ET STRATÉGIE SLM + EDGE AI

- Identifier et prioriser les cas d’usage présentant le meilleur rapport entre valeur métier, performance et coût d’infrastructure.
- Concevoir une plateforme permettant de déployer et administrer des SLM sur plusieurs sites, équipements ou environnements.
- Construire une feuille de route d’industrialisation vers une IA locale, souveraine, scalable et durable.

Maroc Performance

Approche pédagogique

- Support Ecrit et Projection
- Exposés Interactifs, Podcasts et Vidéos
- Brainstorming et Jeux de Rôle
- Mises en Situation pour faciliter l'assimilation
- Cas Pratiques et Labs inclus pour leur impact opérationnel
- Test de Validation des Acquis des Connaissances

Prochaines dates programmées

16 au 25 Nov. 2026	Présentiel - Casablanca
08 au 17 Fév. 2027	Présentiel - Casablanca
05 au 14 Avr. 2027	Présentiel - Casablanca
31 Mai au 09 Juin 2027	Présentiel - Casablanca
26 Juil. au 04 Août 2027	Présentiel - Casablanca

Autres dates possibles sur demande. Contactez-nous pour organiser une session intra-entreprise.

Réservation & Renseignements

Téléphone : +212 522 247 210
Email : contact@maroc-performance.com
Web : https//www.maroc-performance.com

Maroc Performance — Tous droits réservés