Tester et fiabiliser les systèmes d'IA : méthodes, outils et pratiques avancées pour une IA de confiance

2 jours
1 690€ HT

Kereval propose des formations Inter et en Intra-entreprise dispensées par des formateurs experts dans leur domaine. Selon les besoins spécifiques, Kereval est en mesure de dispenser des formations sur-mesure, dont le contenu et la durée sont adaptés à votre contexte.

À mesure que l'intelligence artificielle s'intègre au cœur des produits, des services et des processus métiers, en particulier dans des contextes critiques, la maîtrise des risques liés aux systèmes d'IA devient un enjeu stratégique.

Fiabilité des modèles, qualité des données, robustesse face aux attaques, explicabilité des décisions ou encore conformité réglementaire : tester une IA ne se résume pas aux pratiques du test logiciel classique.

Cette formation propose une approche structurée et opérationnelle du test des systèmes d'IA, des modèles d'apprentissage automatique (ML) aux modèles de langage (LLMs). Elle combine fondements théoriques, méthodes éprouvées et mises en pratique pour garantir une IA de confiance.

Objectifs pédagogiques

À l'issue de la formation, les participants seront capables de :

- Comprendre les spécificités du test logiciel appliqué à l'IA
- Identifier les risques et les principales vulnérabilités de l'IA
- Connaître et mettre en œuvre les méthodes d'évaluation, de test et de vérification pertinentes
- Intégrer les bonnes pratiques de développement en IA et de MLOPS pour une IA fiable

Pré-requis

Connaissances des fondamentaux en IA et du langage de programmation Python

Public formé

- Chefs de projet et responsables techniques
- Ingénieurs et développeurs en IA
- Profils souhaitant renforcer leurs compétences en qualité des systèmes d'IA

Contenu

- Les enjeux de l'IA
  - Qu'est-ce que le test logiciel ?
  - Pourquoi a-t-on besoin de tester l'IA ?
  - Quelle(s) différence(s) avec les logiciels classiques ?
- Les propriétés d'une IA de confiance
  - Qu'est-ce qu'une IA de confiance ?
  - Qualité des données
  - Ethique et équité
  - Explicabilité, interprétabilité
  - Sécurité, sureté et robustesse

Travaux pratiques : mise en œuvre d'outils pour l'équité et l'explicabilité

- Les méthodes et outils pour le test
  - Les étapes de conception des tests
  - Oracles de test
  - Méthodes de test
    - Test métamorphique
      Travaux pratiques : exercices d'identification des relations métamorphiques
    - Attaques des systèmes basés IA
      Travaux pratiques : exercices d'injection de prompts sur LLM
    - Fuzzing ou test à données aléatoires
    - Tests combinatoires
    - Bilan : avantages et inconvénients des méthodes
  - Evaluer la qualité des tests
    - Métriques de couverture
    - Tests de mutation

  Travaux pratiques : mise en œuvre des méthodes de test métamorphique et d'attaques adverses

- Les méthodes de vérification formelle
  - Qu'est-ce qu'une méthode formelle ?
  - Propagation des limites : approximation linéaire et interprétation abstraite
  - Solveurs : SMT et MILP
  - Méthode de branch-and-bound
  - Bilan : comparaison des méthodes

  Démonstration : mise en œuvre de méthodes de vérification formelle

- Développer une IA de confiance
  - Règlementation européenne : AI Act
  - Environnement normatif pour l'IA
  - Bonnes pratiques et MLOps
  - Bonnes pratiques de sécurité

Méthodes pédagogiques

- Formation en présentiel
- Exposé des concepts avec support powerpoint©
- Exercices pratiques : méthodes d'explicabilité, tests métamorphiques, attaques de modèles d'IA, vérification formelle d'un modèle

Modalités d'accès et délais

- Modalités d'inscription : nous contacter
- Délais d'accès : de 1 à 6 mois. Les délais d'accès sont dépendants de la disponibilité de nos formateurs.
- Accessible aux personnes en situation de handicap. Nous contacter pour tout aménagement spécifique.

Modalités d'évaluation

> Le formateur évalue la progression pédagogique du participant tout au long de la formation au moyen de QCM, mises en situation, travaux pratiques.

> Le participant complète également un test de positionnement en amont et en aval pour valider les compétences acquises.

> Pas de certification à la suite de la formation.

Durée

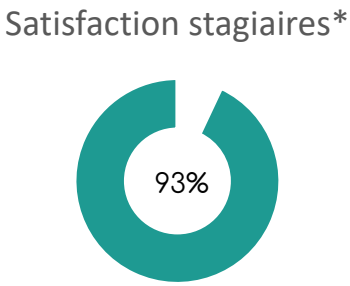
2 journées (14h)

Prix intra – entreprise	Prix inter – entreprise
Nous contacter pour obtenir un devis	1 690 € HT / stagiaire

Nous contacter : contact@kereval.com

Pour toute question y compris les conditions d'accès pour les publics en situation de handicap.

Chiffres clés



* pourcentage calculé sur 1 session de formation en 2024 comprenant 6 stagiaires.