FORMATION · INTELLIGENCE ARTIFICIELLE

NLP – Traitement du langage naturel

Durée de la formation : 3 jours
Tarif : 2 090 € HT par participant
Réf : BI330
Délai d'accès : 2 à 4 semaines
Niveau : Avancé

Prérequis

Bonne pratique de Python. Notions de machine learning ; la formation « Deep Learning – Réseaux de neurones » est un plus.

Public concerné

Data scientists, développeurs, ingénieurs IA

Objectifs pédagogiques

- Connaître les étapes et les techniques du traitement du langage naturel
- Représenter le texte : des sacs de mots aux embeddings
- Utiliser et adapter des modèles transformeurs
- Évaluer un système de NLP et l'intégrer à une application

Objectifs opérationnels

- Préparer et nettoyer un corpus de textes en français
- Classer des textes et extraire des entités
- Utiliser des embeddings pour la recherche sémantique
- Affiner un modèle pré-entraîné avec Hugging Face
- Évaluer la qualité et les biais d'un modèle

Méthodes mobilisées et modalités d'évaluation

Apports théoriques, démonstrations et nombreux travaux pratiques sur des cas concrets. À l'issue de chaque module, le formateur évaluera chacun des participants en fonction des cas pratiques et exercices effectués.

NLP – Traitement du langage naturel

Le traitement du langage naturel (NLP) permet aux machines de comprendre, classer, résumer et générer du texte : analyse de documents, tri de demandes, recherche sémantique, assistants.

Cette formation parcourt les techniques du NLP, des représentations classiques aux transformeurs et aux grands modèles de langage, avec l'écosystème Hugging Face.

Les travaux pratiques portent sur des textes en français, au plus près des cas d'entreprise.

Inscription

Contactez-nous au 01 86 95 71 60 ou à l'adresse formation@adbi.fr pour connaître les dates de notre prochaine session. Inscription au plus tard 10 jours avant le début de la session.

Accessibilité

Nos salles de formation sont accessibles aux personnes à mobilité réduite. N'hésitez pas à nous informer si vous avez des besoins particuliers et s'il y a des besoins d'aménagements spécifiques lors de votre inscription.

Modes de participation

Nos formations sont proposées sous 3 formats :

- Chez nous, près de la place d'Italie (Paris 13e)
- Chez vous, dans vos locaux
- À distance, en classe virtuelle

Attestation

Une attestation de fin de formation sera remise à tous les participants à l'issue de leur parcours.

Programme :

Jour 1

- **Les bases du NLP**
  - Cas d'usage en entreprise
  - Prétraitement : segmentation, normalisation, lemmatisation
  - Représentations classiques : sacs de mots, TF-IDF
- **Premiers modèles**
  - Classification de textes
  - Reconnaissance d'entités nommées
  - Travaux pratiques : tri automatique de demandes

Jour 2

- **Représentations modernes**
  - Embeddings de mots et de phrases
  - Recherche sémantique et similarité
- **Transformeurs**
  - Attention et architecture des transformeurs
  - Modèles encodeurs et décodeurs
  - Écosystème Hugging Face : modèles, jeux de données, pipelines
  - Affiner un modèle pré-entraîné

Jour 3

- **Grands modèles de langage**
  - Génération de texte, résumé, traduction
  - Prompting et RAG : principes
  - Choisir entre modèle spécialisé et grand modèle généraliste
- **Évaluer et déployer**
  - Métriques de qualité et jeux d'évaluation
  - Biais, confidentialité des données et sobriété
  - Servir un modèle sous forme d'API
- **Atelier de synthèse**
  - Analyse automatique d'un corpus de documents