IA : Traitement Automatique du Langage et Applications Modernes

Lien : https://innov-systems.com/formation/ia-traitement-automatique-du-langage-et-applications-modernes

DURÉE
3 jours (21h)

RÉFÉRENCE
BSI534

CATÉGORIE
Data Science et Data Engineering

Innov Systems

OBJECTIFS DE LA FORMATION

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

- Comprendre en profondeur les principes du traitement automatique du langage (TAL/NLP)
- Savoir mettre en œuvre les techniques de base et avancées avec Python et bibliothèques modernes (NLTK, spaCy, Transformers)
- Être capable d’appliquer le NLP à des cas concrets : classification, traduction, analyse de sentiments, résumé automatique
- Découvrir les enjeux et limites actuelles du NLP (biais, qualité des données, scalabilité)

POUR QUI ?

- Ingénieurs
- Analystes
- Data Scientists
- Data Analysts
- Data Stewards
- Développeurs
- Chefs de projet IT/Data

Innov Systems

Programme détaillé

1 / INTRODUCTION AU TRAITEMENT AUTOMATIQUE DU LANGAGE (NLP)

- Compréhension des objectifs du NLP : analyse, génération et compréhension du langage naturel
- Les étapes clés : tokenisation, nettoyage, lemmatisation, reconnaissance d'entités nommées
- Historique : approche statistique, machine learning, deep learning
- Panorama des outils : NLTK, spaCy, Hugging Face Transformers, Watson NLU
- Cas d'usage : détection de spam, traduction, assistants virtuels, chatbots, moteurs de recherche, veille et analyse d'opinions
- Atelier pratique : exploration de corpus textuels et premiers traitements simples avec Python

2 / MANIPULATION DE TEXTES AVEC PYTHON

- Installation et prise en main des packages NLP (NLTK, spaCy)
- Importation de corpus et jeux de données textuels
- Tokenisation en mots et phrases
- Nettoyage et normalisation des textes (stopwords, ponctuation, mise en minuscule, accents)
- Techniques de stemming et lemmatisation
- Étiquetage morpho-syntaxique (POS tagging)
- Atelier : préparation d'un corpus en vue d'analyses ultérieures

3 / ANALYSES DE TEXTES ET RECONNAISSANCE D'ENTITÉS

- Analyse de fréquence et représentation des textes (n-grams, distributions de termes)
- Visualisation : concordances, nuages de mots, dispersion des termes
- Reconnaissance d'entités nommées (NER) avec NLTK et spaCy
- Extraction d'informations pertinentes dans des textes volumineux
- Atelier : analyse de corpus et extraction de mots-clés/caractéristiques

4 / ANALYSE DE SENTIMENTS ET CLASSIFICATION

Innov Systems

- Présentation des méthodes : règles, lexiques, machine learning, deep learning
- Utilisation de nltk.sentiment et spaCy pour l’analyse de sentiments
- Atelier : mise en œuvre d’une analyse de sentiments sur des avis clients
- Intégration de scikit-learn : vectorisation (Bag of Words, TF-IDF), classification supervisée
- Atelier : entraînement d’un modèle de classification textuelle avec scikit-learn

5 / INTRODUCTION AUX MODÈLES AVANCÉS ET TRANSFORMERS

- Limites des approches classiques et apports des modèles modernes
- Présentation de Word2Vec, GloVe, embeddings contextuels (ELMo, BERT)
- Découverte des Transformers et de Hugging Face
- Atelier : utilisation d’un modèle pré-entraîné pour la classification ou la traduction automatique

6 / ÉTUDES DE CAS CONCRETS

- Étude de cas 1 : construction d’un chatbot simple basé sur NLP
- Étude de cas 2 : résumé automatique d’articles ou de documents
- Étude de cas 3 : extraction d’opinions dans des tweets
- Atelier final : projet de synthèse avec choix libre parmi plusieurs cas pratiques

7 / ENJEUX ET BONNES PRATIQUES

- Qualité des données textuelles : biais, ambiguïtés et bruit
- Scalabilité : traitement de corpus volumineux
- Bonnes pratiques pour déployer une solution NLP en production
- Ouverture : applications émergentes (traduction neuronale, agents conversationnels génératifs, recherche sémantique)

Innov Systems

Approche pédagogique

- Support Ecrit et Projection
- Exposés Interactifs, Podcasts et Vidéos
- Brainstorming et Jeux de Rôle
- Cas Pratiques et Labs inclus pour leur impact opérationnel
- Test de Validation des Acquis des Connaissances

Prochaines dates programmées

13 au 15 Jan. 2027	Présentiel - Casablanca
10 au 12 Mars 2027	Présentiel - Casablanca
05 au 07 Mai 2027	Présentiel - Casablanca
30 Juin au 02 Juil. 2027	Présentiel - Casablanca
01 au 03 Sep. 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@innov-systems.com
Web : https//www.innov-systems.com

Innov Systems — Tous droits réservés