

Joseph Beasse

Ingénieur en Intelligence Artificielle et Sciences Cognitives

Cherche un CDI en data science, IA et/ou NLP pour Septembre 2024

✉ jbeasse.pro@gmail.com

🐙 github.com/jbeasse

🌐 linkedin.com/in/josephbeasse

+33 (0)6 95 57 12 48

FORMATION

•Diplôme d'ingénieur en Intelligence Artificielle et Sciences Cognitives

2021-24

ENSC & ENSEIRB-MATMECA, Bordeaux

Note: 16/20

Cours pertinents : Algorithmes de Recherche, IA pour les Jeux Vidéo, Apprentissage automatique, profond et par renforcement, Vision par ordinateur, Intelligence artificielle musicale, Ingénierie logicielle pour l'IA, NLP.

•Classes Préparatoires Aux Grandes Écoles (Spécialité Physique & Sciences de l'Ingénieur)

2019-2021

Chateaubriand, Rennes

EXPÉRIENCES

•Stagiaire ingénieur IA et NLP

Février - Septembre 2024

Orange

Lannion

- Raffinement du code existant et réalisation d'analyses statistiques et de corrélation sur les données d'alarme.
- Implémentation du transfert d'apprentissage pour le modèle de masquage de langage non supervisé (BERT modifié).
- Automatisation de l'extraction des causes potentielles des alarmes à partir des manuels des fournisseurs en utilisant divers LLM quantifiés, des modèle d'embeddings 'sentence-transformers' et des algorithmes de clustering.
- Identification des causes racines des alarmes pour accélérer l'intervention des techniciens.

•Stagiaire ingénieur IA et data scientist

Avril - Septembre 2023

Sopra Steria

Toulouse

- Développement d'outils basés sur l'IA pour simplifier la rédaction et la vérification des appels d'offres (RFP).
- Création d'outils de correction et de relecture d'appels d'offres basés sur le traitement du langage naturel.
- Synthétisation et reformulation de textes et documents pour améliorer leur clarté et cohérence.
- Indexation et recherche avancée de documents avec Elasticsearch.
- Évaluation de la similarité entre documents pour identifier des contenus redondants ou complémentaires.

PROJETS

•Approche de réduction d'échelle pour l'apprentissage contrastif : SimCLR

Adaptation du framework d'apprentissage contrastif auto-supervisé SimCLR pour des environnements à ressources limitées.

- Investigation des effets des augmentations de données et taille de batch sur les capacités d'apprentissage du modèle.
- Utilisation des Vision Transformers (ViT) pour l'extraction des caractéristiques (ultérieurement ResNet18).
- Stack technique : PyTorch, CIFAR-10, ResNet18, Vision Transformers.

•Génération Automatisée de Graphes de Connaissance

Extraction et visualisation de relations textuelles à partir de fichiers PDF pour générer un graphe de connaissances.

- Extraction de triplés 'sujet-prédicat-objet' de relations décrites dans des documents économiques à l'aide d'un modèle de langage: REBEL (Relation Extraction By End-to-end Language generation).
- Construction de graphes de connaissances avec Pygraft et rdflib.
- Interface utilisateur basée sur Streamlit pour la visualisation des graphes générés.
- Stack technique : Python, nltk, llm, pygraft, rdflib, torch, transformers, Streamlit, Flask.

•DeepLanguage : Identification Automatique de Langues

Développement d'une IA capable d'identifier les langues parlées dans des fichiers audio.

- Collecte et pré-traitement des données audio pour les convertir en format wav et les préparer pour l'apprentissage.
- Extraction des caractéristiques en utilisant des spectrogrammes et des coefficients cepstraux de Mel (MFCCs).
- Développement et évaluation de modèles CNN et RNN pour l'identification des langues.
- Création d'un site web avec Flask pour permettre l'utilisation de l'algorithme via une interface utilisateur.
- Stack technique : Python, TensorFlow, Librosa, Flask, HTML, CSS, JavaScript.

COMPÉTENCES TECHNIQUES ET INTÉRÊTS PERSONNELS

Langages: Python, R, SQL

Frameworks: Pytorch, Tensorflow, Flask, FastAPI

Bibliothèques: Transformers, Nltk, Pandas, Numpy, Scikit-learn

Outils de Développement: Linux, Git, Azure

Domaines d'Intérêt: Finance, Recherche en IA et NLP, Développement personnel

Certifications: DeepLearning.AI: Tensorflow Professional (03/2023) | Generative Deep Learning (03/2023)

Langues: Français (natif), Anglais (C1: IELTS 7.5), Espagnol (B2), Allemand, Russe