

## PROFIL



Data Scientist avec 6 ans d'expérience dans le développement de modèles de Machine Learning avancés (Deep Learning, NLP-LLM, RL) pour des applications innovantes dans les secteurs des télécommunications, de la vente au détail et de l'énergie.

## COMPETENCES TECHNIQUES



- **Développement web:** Java J2EE, NextJS, Flask, Dash-Plotly, Gradio
- **Langages de programmation :** Python, SQL, Java, JavaScript
- **Bases de données :** MySQL, Postgres, Neo4j
- **Méthodologies et formalismes :** Agile (Scrum), UML, Merise
- **Frameworks de deep learning:** TensorFlow, PyTorch, PyTorch-Lightning, Keras
- **Data Science:** Scikit-learn, Pandas, NumPy, SciPy, Seaborn, Matplotlib, Statsmodels
- **Data Engineering:** Spark, Kafka, Airflow, Docker, Azure

## FORMATIONS



- **2021 : Doctorat en Intelligence Artificielle** - Sorbonne Université, Institut des Systèmes Intelligents et de Robotique
- **2016 : Master en Data Science** - Université Paris 13
- **2014 : Diplôme d'Ingénieur en Informatique** - Institut National de Statistique et d'Économie Appliquée (INSEA)
- **2010 : Classes Préparatoires** aux Grandes Écoles (MPSI/MP)

## CERTIFICATIONS



- Fundamentals of Reinforcement Learning- ALBERTA UNIVERSITY.
- Mathematics for Machine Learning- Imperial College London.
- Building Batch Data Pipelines on Google Cloud- Google.
- Google Cloud Big Data and Machine Learning- Google.
- Deep Learning Specialization- Deep learning.AI.
- MLOPS course - FDSL (full stack deep learning).
- Andrej Karpathy courses on NLP.



## LANGUES

- Anglais : courant.
- Français : courant.



Akkodis  
NLP-Lead

Novembre 2024

> **Projet : EU-DREAM**

> **Equipe: (1 PM, 2 Data Engineer, DevOps, 1 BA)**

> **Contexte** : Le projet EU-DREAM vise à accélérer l'innovation numérique dans le secteur de l'énergie. En tant que responsable NLP, j'ai développé un intermédiaire linguistique pour rendre les informations complexes du marché de l'énergie accessibles à tous. Cet outil, basé sur l'IA, traduit le jargon technique en langage clair. J'ai également contribué à la conception de l'architecture de la plateforme de données, permettant la création de Jumeaux Numériques en temps réel pour optimiser la gestion de l'énergie au niveau des ménages et des communautés.

- **Conception et développement d'un system agentique (ReAct) en utilisant Langgraph.**
- **Conception et implémentation d'un modèle de génération de réponses (RAG) auto-réflexif.**
- **Développement d'outils d'interaction pour l'agent (API pricing, météo).**
- **Développement d'une interface utilisateur pour le traitement du langage naturel (NLP) intermédiaire à l'aide de Gradio.**
- **Conception d'un modèle de préférence d'utilisateur en se basant sur la mémoire de l'agent.**
- 

> **Environnement Technique :**

- o **Agent-LLM, NLP, Langgraph, Gitlab, Python, Gradio, LLM, Langchain, Postgres Pg, Azure-OpenAI, FASTAPI.**

---

Akkodis

Août 2024 - Novembre 2024

Senior Data Scientist

> **Projet : GPT-Adecco**

> **Equipe : (1 PM, 3 Développeurs, 4 Data Scientist)**

> **Interlocuteurs externes : Adecco.**

> **Contexte** : Chez Adecco, j'ai contribué à l'initiative Non-Discrimination, visant à détecter et éliminer les termes discriminants dans le système GOELAND pour garantir l'équité dans les décisions. J'ai également développé un chatbot documentaire, permettant de créer des assistants métier pour améliorer l'accès à l'information et l'interaction utilisateur. Ces projets reflètent mon engagement à promouvoir l'inclusion et l'efficacité à travers des solutions technologiques.

- **Parsing et indexation de divers documents (PDF, Word, PPTX et Excel) sur Azure Vector Search.**
- **Implémentation d'un RAG sur des données internes de recrutement et juridiques avec LangChain.**
- **Amélioration de la qualité du retrieval afin d'obtenir uniquement des ressources pertinentes.**
- **Développement d'une API de classification des commentaires discriminatoires saisis par les RH, en utilisant du prompt engineering avec GPT-4.**

➤ **Environnement Technique :**

- Azure Devops, Azure ML, Streamlit, Pandas, Hugging-Face, Weight&Bias, Python, Gradio, Chatbot, LLM, Langchain, Azure AI search, Azure-OpenAI, Azure Document-intelligence, NextJS.
- 

**Akkodis Research**

**Janvier 2024 – Juillet 2024**

Data Scientist/Manager en innovation

➤ **Projet : TAIE : TalentAI Enhance**

➤ **Equipe : (1 PO, 3 Développeurs, 3 Data Scientist)**

➤ **Interlocuteurs externes : Adecco.**

➤ **Contexte** Le projet TAIE vise à optimiser l'efficacité des processus RH tout en améliorant l'expérience des employés. Il se concentre sur cinq axes : la personnalisation des processus RH grâce à l'IA, l'analyse contextuelle des données RH via le NLP, l'interopérabilité des systèmes RH en garantissant la sécurité des données, l'évaluation éthique des décisions prises par IA, et le développement de modèles prédictifs pour orienter les parcours de carrière selon les tendances du marché et les aspirations individuelles.

- **Développement d'un simulateur de salaire permettant de prédire une fourchette salariale à partir du profil des candidats, en utilisant PyTorch Lightning sur des données fournies par Adecco.**
- **Mise en place d'un système RAG (Retrieval-Augmented Generation) pour accéder et analyser les dossiers de compétences des collaborateurs avec LangChain.**
- **Développement d'une application de gestion de ressources humaines intelligente intégrant le Simulateur et RAG.**

➤ **Environnement Technique :**

- Azure Devops, Azure ML, Streamlit, Pandas, Pytorch, Pytorch-Lightning, Weight&Bias, Python, Gradio, Chatbot, LLM, Langchain, Azure AI search, Azure-OpenAI.
- 

**Akkodis Research**

**Janvier 2023 – Décembre 2023**

Data Scientist/Manager en innovation

➤ **Projet : DAIN : Data Analytics for Intelligent Networks**

➤ **Equipe : (5 Développeurs, 1 QA, 1 BA)**

➤ **Interlocuteurs internes : Centre de service Akkodis**

➤ **Interlocuteurs externes : KIABI.**

➤ **Contexte** : Le projet DAIN au sein d'AKR vise à développer un outil prédictif pour les centres d'appels afin d'anticiper la charge de travail, notamment le volume d'appels entrants, et d'automatiser la planification du personnel. Ce projet s'inscrit dans un contexte de croissance des réseaux intelligents, avec l'essor de technologies telles que l'IoT et la 5G. L'objectif est d'améliorer l'allocation des ressources humaines, optimiser les coûts d'exploitation, et offrir une meilleure qualité de service grâce à des prévisions précises et des systèmes de planification intelligents basés sur l'apprentissage par renforcement.

- **Définition de la stratégie et des objectifs du projet via une roadmap.**
- **Réalisation d'une veille technologique et scientifique pour prévenir les risques d'obsolescence technologique.**
- **Analyse des besoins métiers et techniques, rédaction des spécifications fonctionnelles sous forme de User Stories, et priorisation du Backlog.**
- **Création d'un pipeline d'ingestion, traitement et stockage de données sous forme de séries temporelles (appel, mail, SMS, promotion, vente privée, etc.).**
- **Elaboration des différents Dashboard de suivi du centre de service avec Dash-plotly.**

- Entraînement et comparaison de différent modèle de prévision (ARIMA, Prophete, LSTM, DeepAR) pour prédire le flux entrant (ex. nombre d'appel entrant par minute).
- Automatisation de la planification du personnel en combinant le modèle SARSA (Reinforcement Learning) et ERLANG-C afin de maintenir un bon niveau de service (appels décrochés en moins de 20 s > 80%).
- Orchestration des différents modules de l'application avec Airflow (DAGS).
- Développement d'une application web avec Flask intégrant les modules et dashboards.
- Suivi des résolutions des bugs.
- Déploiement de l'application sur Azure.
- Ecriture de la fiche CIR (crédit impôt recherche).
- 

➤ Environnement Technique :

- JIRA, Flask, Dash-Plotly, docker, Tensorflow, OpenAI gym, Gitlab, Mysql, Python, Pandas, SQL, Airflow et MLflow.

## Akkodis Reasearch

Mai 2022 –Décembre 2022

Data Scientist

➤ Durée : 8 mois

➤ Projet : Création d'une application pour la détection des anomalies dans les CDR (call detail records)

➤ Interlocuteurs externes : Orange

- Réalisation d'une veille technologique et scientifique pour atténuer les risques d'obsolescence technologique.
- Développement d'une nouvelle approche dynamique de détection d'anomalie en temps réel en se basant sur des réseaux de neurones récurrent et convolutif.
- Elaboration d'un Dashboard pour visualiser les données CDR et les anomalies détectées pour chaque cellule télécom.
- Amélioration de la performance du modèle en exploitant des données auxiliaires : twitter, météo et précipitation.
- Ecriture de la fiche CIR.

➤ Environnement Technique: Tensorflow, Dash-Plotly, Spark, MySQL, Python, Folium, Seaborn, Flask, Scikit-Learn, Pandas

## Université Paris 8

Mai 2021 –Décembre 2021

Attaché Temporaire d'Enseignement et de Recherche

➤ Durée : 8 mois

➤ Projet : Enseignement

- Enseignement : Systèmes de gestion de bases de données relationnelles (Master).
- Enseignement : Programmation orientée objet avec Java (Master).
- Enseignement : Introduction à la programmation (Licence).
- Co-supervision d'un mémoire de Master.

➤ Environnement Technique : Java, MySQL, Python.

➤ **Durée : 3 ans et 3 mois**

➤ **Projet : ANR-Social Touch (Project-ANR-17-CE33-0006)**

➤ **Equipe : (LTCI-Télécom ParisTech, ISIR, CNRS, Heudiasyc)**

➤ **Contexte :** Ce projet étudie l'intégration du sens du toucher dans les systèmes interactifs pour améliorer la communication et l'émotion entre humains et machines. Il explore le rôle social du toucher pour favoriser l'engagement, en développant de nouvelles techniques d'interaction homme-machine et en évaluant l'acceptabilité du toucher dans les communications interpersonnelles médiées et avec des agents conversationnels, comparant les interactions humaines et virtuelles.

- Développement d'un modèle Deep Learning pour la reconnaissance de 14 gestes tactiles sur un robot zoomorphique, avec des résultats de pointe sur les corpus CoST et Haart.
- Développement d'un nouveau Protocol expérimental basé sur l'imitation pour la collecte des données tactile socio-affectives dans une interaction Humain-Machine.
- Développement d'un humanoïde à peau artificielle piézorésistive pour la perception tactile.
- Développement d'un modèle personnalisé pour la reconnaissance automatique des émotions et des traits individuels (personnalité et empathie) à partir de signaux tactiles sociaux.
- Coordination active avec les équipes partenaires pour garantir l'intégration harmonieuse de nos livrables.
- Présentation de mes recherches lors de conférences internationales : ACII (Cambridge) et ICMI (Utrecht).

➤ **Environnement Technique :**

- Tensorflow, Transformers, Arduino, Python, Scikit-Learn, Psychopy, Inference statistiques, pandas, seaborn, matplotlib, scipy, statsmodel, Test d'hypotheses, Quatrics forms.

➤ **Durée : 10 mois**

➤ **Projet : Contribution au développement de SyncPy, une bibliothèque open source pour l'analyse de la synchronie dans les interactions humain-humain et humain-machine.**

➤ **Equipe : (1 PU, 1 PostDoc, 1 Ingénieur d'étude, 2 développeurs)**

- Développement d'un modèle de détection automatique de l'imitation en utilisant le modèle BOW (bag of words) sur des features HOG et HOF pour caractériser les gestes des participants et le one-class SVM pour mesurer la similarité entre ces gestes.
- Développement d'un modèle de détection automatique de l'imitation avec un modèle RNN-LSTM (Long short-term memory) et des Deep features (Convolutional neural network features: Google InceptionV3).
- Mise en place d'une expérience pour la collecte de données de synchronie à l'aide d'un optiTrack
- Implémentation d'un modèle de réseaux de neurones siamois pour la mesure de similarité de geste sur la base de données Weizmann (Action recognition DB)
- Evaluation des modèles à l'aide d'une matrice de cross-corrélation

➤ **Environnement Technique :**

- Python, Scikit-Learn, Tensorflow, FFMPEG, OpenCV, OptiTrack, Pandas, statsmodel, Numpy, Scipy, Matplotlib, Seaborn.

- **Durée : 7 mois**
  - **Projet : Création d'un Convertisseur de code sémantique pour les applications mobiles**
  - **Equipe : CEO, CTO, CMO, 1 architecte, 1 PM, 20 développeurs**
    - **Transformation et projection du code source des applications iOS vers Android en se basant sur le Reverse Engineering et MDE (model Driven Engineering).**
    - **Amélioration du noyau de l'application pour assurer une transitivité entre différentes Plateformes (iOS, Android, Windows Phone)**
  - **Environnement Technique :**
    - **Java, EMF, Karaf, Neo4j, Jira, Redmine, kdm, astm, iOS, Android, STG, Xpand, kdm, astm.**
-