

TADIELLO SÉBASTIEN

Machine Learning Engineer

DOSSIER DE COMPÉTENCES

PRÉSENTATION

Je suis Sébastien Tadiello, un Machine Learning Engineer et MLOps, avec une expertise marquée en apprentissage supervisé et non supervisé ainsi que dans le déploiement des solutions développées (API RESTful, Web App) ou dans l'incorporation des solutions développées pour des structures backend préexistantes.. Ce dossier vise à détailler mes compétences techniques, expériences professionnelles, et contributions académiques.

EXPÉRIENCE PROFESSIONNELLE

ARTIFICIAL INTELLIGENCE ENGINEER

Groupe Pacte Novation Paris (Janvier 2024)

- Objectif:
 - Recherche et développement sur les modèles de langage de grande taille (LLM) : exploration des avancées technologiques, évaluation de nouveaux modèles, et mise en œuvre de solutions innovantes pour améliorer les performances et l'intégration des LLM dans les applications d'entreprise.
- Réalisation(s) :
 - Mise en place d'une infrastructure technique, permettant le développement de solutions d'intelligence artificielle.
 - Développement et déploiement d'un chatbot interne : Création d'un chatbot utilisant différents modèles de LLM open source.
 - Optimisation de la mémoire du chatbot : Amélioration de la gestion de la mémoire pour optimiser les performances du chatbot et sa capacité à suivre une conversation structurée.
 - Mise en place d'un système de récupération d'informations (RAG) : Implémentation d'un RAG permettant d'accéder et de traiter les documents internes de l'entreprise.

MACHINE LEARNING ENGINEER

SCOR SE Paris (Septembre 2022 - Décembre 2023)

- Objectifs:
 - Identifier une problématique auprès d'un commanditaire.
 - Identifier les solutions techniques étant les plus à même de répondre aux problèmes et concevoir une solution digitale sur la base des avancées en Data et IA.
 - Développer et mettre en production la solution.
- Réalisation(s) :
 - Entretiens clients, et identification des problématiques métiers.
 - Développement et optimisation de pipelines de données avancés, garantissant une intégration et un traitement efficace des informations dans un écosystème Azure en utilisant des solutions comme Databricks et Azure DevOps .
 - Traitement automatique du langage naturel (NLP), avec des modèles de pointe tels que BERT, GPT et des techniques de reconnaissance d'entités nommées (NER) pour analyser et extraire des informations pertinentes de vastes ensembles de données textuelles. Mise en place d'API (FastAPI) pour rendre le service accessible.
 - Élaboration d'algorithmes de correspondance de données, utilisant la similarité cosinus et la distance de Levenshtein, pour associer automatiquement des listes de clients provenant de différents services, améliorant ainsi significativement l'efficacité des processus internes.
 - Développement et mise en production de modèles de machine learning visant à classer automatiquement des documents internes de l'entreprise permettant de fluidifier leur traitement par les équipes de data capture. Déploiement des algorithmes sur un modèle CI/CD en utilisant des outils d'automatisation tels que GitLab et DockerCompose permettant une amélioration des processus sans interruption des services.
 - Déploiement des tableaux de bord interactifs via Tableau et Power BI, offrant des insights data-driven et facilitant la prise de décision stratégique au sein de l'organisation.
 - Extraction d'information web avec la création d'algorithmes de Scraping web.
- Environnement(s) technique(s) :
 - Kubernetes, Azure Databricks, SQL, Python, Docker, DockerCompose, Bash, Git, TensorFlow, Tableau, PowerBi

DATA SCIENTIST

Recherche à l'UMRS 1158 Inserm-Sorbonne Université (Octobre 2021 - Juin 2022)

- Objectif:
 - Mise en place d'un protocole d'étude expérimentale portant sur la modulation de la dyspnée dans l'interaction humain / robot.
- Réalisation(s) :
 - Analyses de données de langage naturel (NLP) afin d'identifier les émotions des sujets lors d'interactions avec un robot.
 - Étude de la synchronie inter-individuelle en se basant sur l'analyse de signaux biologiques comme ECG, mesure du flux respiratoire et les temps de fixation visuelle.
 - Développement d'algorithmes pour l'interaction humain-robot (suivi de visage, mouvement synchrone, etc).
- Environnement(s) technique(s) :
 - SQL, Python, Micro Python, R, Bash, Git, TensorFlow

DATA SCIENTIST

Université de Bordeaux, Institut de Neurosciences Cognitives et Intégratives d'Aquitaine (INCIA) (février 2021 - juin 2021)

- Objectif:
 - Analyse statistique et présentation des résultats aux équipes de recherche.
- Réalisation(s) :
 - Analyse et pre-processing des données. Analyse statistique des données visant à les rendre intelligibles auprès des équipes de recherche.
 - Mise en place de modèle prédictif (random forest).
 - Présentation des résultats de la recherche lors du congrès 'COLlectif pour la Recherche Transdisciplinaire sur les Interfaces Cerveau-Ordinateur' (Cortico).
- Environnement(s) technique(s) :
 - Python, R, Bash, Git, TensorFlow

FORMATIONS

MASTER DATA ET INTELLIGENCE ARTIFICIELLE

Hetic Paris (2021-2022)

- Python | R | SQL
- API Docker & Cloud
- Machine Learning | Réseaux de Neurones
- Big Data Architecture

MASTER DATA SCIENCE

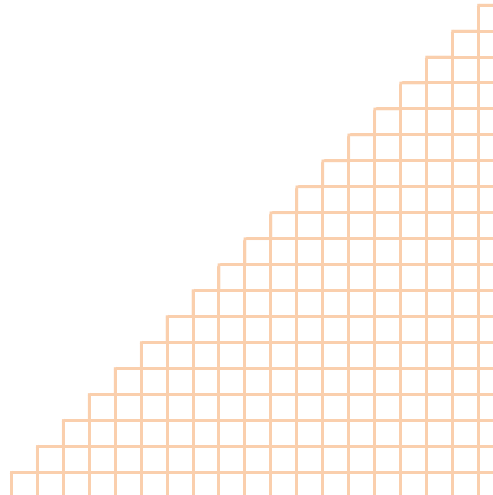
Université Gustave Eiffel à Paris (2021-2022)

Enseignements en data science dispensés à l'ESIEE (Ecole Supérieure d'Ingenieur en Electrotechnique et Electronique):

- Machine learning
- Text mining
- Analyse de réseaux
- Statistiques, algèbre linéaire, optimisation

MASTER DYNAMIQUE COGNITIVE ET SOCIOCOGNITIVE

Université Paul-Valéry Montpellier (2019-2021)

- Compréhension des Mécanismes Cognitifs
 - Modélisation Computationnelle des Fonctions Cognitives
 - Projet de Recherche ou Applications Pratiques
 - Optimisation des facteurs humains
 - Méthode Scientifique
- 

COMPÉTENCES TECHNIQUES

- **Langages de Programmation :**
 - Python
 - R
 - SQL
 - C++
 - Bash
 - Java
- **Machine Learning & Deep Learning :**
 - TensorFlow
 - Scikit-Learn
 - Keras
 - PyTorch
- **MLOps:**
 - Docker
 - FastAPI
 - DockerCompose
 - Kubernetes
 - Vercel
 - GitHub Action
 - Flask / Django / Streamlit
- **Versioning :**
 - Git
 - MLflow
 - DVC

CONTACT



✉ sebastientadiello@gmail.com

☎ 07 60 15 96 34

in <https://www.linkedin.com/in/sebastien-t-ababa2128/>

🐙 <https://github.com/sebDtSci>