

# Giovanni Gozzo

DATA - SCIENTIST



## CONTACT



780 Avenue Georges Freche  
34170, Castelnau-Le-Lez



06 22 39 37 98



gio.gozzo84@gmail.com

## COMPÉTENCES

- Création d'API
- Web scrapping
- Requêtage de Bases de Données Massives (Big Data)
- Machine Learning ( SVM, LR, RF)
- Vectorisation (TF-IDF, BOW)
- Clustering ( K-Means, DB-SCAN)
- HuggingFace ( LLM , Embedding)

## LANGUAGES

- **Python:** (LangChain, Torch, SkLearn, Transformers, FastAPI, Conda, HuggingFace, Pandas)
- **Java**
- **SQL** (Oracle, MySQL, PostGres, ChromaDB, Qdrant )
- **PHP** ( Symfony )

## LANGUES

- Français ( Courant )
- Anglais ( B2 )

## RECHERCHE MASTER DATA-SCIENCE

Étudiant en BUT Informatique passionné par la data science, je souhaite poursuivre ma formation dans un Master en Science des Données. J'ambitionne de développer davantage mes connaissances pour résoudre encore des problématiques complexes et contribuer à des projets d'IA innovants.

## FORMATION

2025	BUT (Bachelor universitaire et technologique) Informatique
2022	IUT Montpellier-Sète / Montpellier
2022	DCG (Diplôme Comptabilité et Gestion)
2021	Lycée Théodore Aubanel / Avignon
2021	Baccalauréat Mention assez bien / Spécialité Mathématiques, SES , NSI
2018	Lycée Théodore Aubanel / Avignon

## EXPÉRIENCE PROFESSIONNELLE

PRÉSENT	Alternant - Data-Scientist
-	ABES - Montpellier
SEPTEMBRE 2024	<ul style="list-style-type: none"><li>• Développement de modèles de machine learning (KNN, Régression Logistique, SVM, Random Forest).</li><li>• Utilisation du clustering avec réduction de dimensionnalité (LDA, PCA) pour la visualisation des clusters.</li><li>• Requêtage de données massives (plusieurs Go)</li><li>• Vectorisation de données textuelle ( Embedding )</li><li>• Déploiement autonome des pipelines et modèles sur un serveur OVH.</li></ul>
JUIN 2024	Stagiaire - Data-Scientist
-	YOOZ - Aimargues
AVRIL 2024	<ul style="list-style-type: none"><li>• Utilisation LLM et d'embeddings pour automatiser des tâches d'analyse textuelle.</li><li>• Développement de pipelines de traitement des données textuelles (prétraitement, nettoyage, optimisation des embeddings).</li><li>• Déploiement autonome de solutions de NLP sur des serveurs Azure.</li></ul>