

Naveen JOHNSON VALLAVANATT

Nice, France ▪ E-mail : naveenjohnsonv@gmail.com ▪ Tél : +33 6 02 56 18 12 ▪ LinkedIn : in/naveenjv ▪ Portfolio : naveenjohnsonv.fr

PROFIL

Data Scientist avec Master en Data Science (EURECOM) et 2+ ans d'expérience en analytics. Expertise en Python, PyTorch, TensorFlow, NLP, deep learning, GenAI et science des données. Recherche poste Data Scientist ; ouvert aux postes Data Engineer/Analyst.

EXPÉRIENCE PROFESSIONNELLE

Stagiaire Recherche Data Science - Planification IA / NLP

- ALTEN Labs
- Mars 2025 - Août 2025, Valbonne, Alpes-Maritimes, France
- Développement de pipeline LLM générant spécifications PDDL formelles exécutées via planificateur Fast Downward pour récits interactifs jouables
 - Amélioration de 60% de fiabilité via validation automatisée schéma/cohérence et mise en cache d'état pour auditabilité

Consultant Associate - SAP Analytics

- Blueprint Technologies Pvt Ltd
- Décembre 2021 - Août 2023, Bengaluru, Karnataka, Inde
- Livraison de 20+ tableaux de bord SAC avec analyses statistiques détaillées pour parties prenantes pharma/agrochimie
 - Construction de migrations ETL entre SAP ECC et BW/4HANA ; conception de logique KPI et pilotage UAT sur 5+ projets

Stagiaire - SAP Analytics

- Blueprint Technologies Pvt Ltd
- Juillet 2021 - Décembre 2021, Bengaluru, Karnataka, Inde
- 200+ heures de formation sur SAC, BW, HANA, Python, SQL; contribution aux prototypes et à la documentation ETL/UAT
 - Automatisation des contrôles de qualité des données en utilisant Python/Pandas et des extractions SQL

FORMATION

Diplôme Master en Informatique, mention Data Science

EURECOM ▪ Biot, France

2023 - 2025

- Sujet principales : Machine Learning, Reinforcement/Deep Learning, Inférence Statistique Avancée, Implémentation SGBD

Diplôme en Data Science

Indian Institute of Technology, Madras ▪ Chennai, Tamil Nadu, Inde

2021 - 2024

Diplôme d'Ingénieur (B.Tech) en Électronique et Télécommunications

National Institute of Technology, Calicut ▪ Kozhikode, Kerala, Inde

2016 - 2020

PROJETS

- Ingénierie des Connaissances à l'Ère des LLM
- EURECOM ▪ 2024
- Reproduction/extension d'expériences Text2KGBench comparant LLMs de pointe pour extraction triplets RDF avec évaluation précision/rappel/F1
 - Développement de pipeline LLM conscient ontologies : GPT-4o atteignant AUC 0.89 vs Qwen2.5 32B ; intégration dataset DBpedia-WebNLG, Wikidata-TekGen, Odeuropa avec ontologies CIDOC-CRM

Analyse de Sentiments avec BERT

- EURECOM ▪ 2024
- Classification de tweets basée sur BERT atteignant 78% de F1-score ; surpassant approches NLP traditionnelles (TextBlob, VADER) via preprocessing complet

COMPÉTENCES

Technique : Python, SQL, R, PyTorch, TensorFlow, XGBoost, CNN, RNN, transformers, BERT, GPT-4o, HuggingFace, LangChain

Data/Cloud : Spark, ETL, MongoDB, Elasticsearch, GCP (Vertex AI, BigQuery), AWS, Docker, Git, CI/CD, SAP SAC, Power BI

Spécialisé : NLP, GenAI (prompt engineering, RAG, fine-tuning), séries temporelles, computer vision, graphes connaissances

Langues : Anglais - Courant (C2), Français - Intermédiaire (B1)