

Naveen JOHNSON VALLAVANATT

📍 Nice, France ✉ naveenjohnsonv@gmail.com ☎ +33 6 02 56 18 12 🌐 in/naveenjv 🌐 naveenjohnson.fr

PROFIL

Data Scientist titulaire d'un Master en Data Science (EURECOM) avec 2+ ans d'expérience en tant qu'un Business Analyst. Maîtrise de Python, PyTorch, TensorFlow, NLP, deep learning et ingénierie des connaissances. Recherche poste Junior 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

- Développé un pipeline LLM pour générer des spécifications de tâches formelles avec planificateur open-source
- Réduit les échecs d'exécution de 60% via contrôles de schéma/cohérence avec corrections automatisées
- Implémenté hachage d'état et mise en cache pour traçabilité et reproductibilité des plans

Consultant Associate - SAP Analytics

Blueprint Technologies Pvt Ltd

Décembre 2021 - Août 2023, Bengaluru, Karnataka, Inde

- Livré 20+ tableaux de bord SAC avec analyses et vues d'écarts pour les parties prenantes pharma et agrochimie
- Construit des migrations ETL entre SAP ECC et BW/4HANA (ADSOs, CompositeProviders, transformations)
- Conçu logique KPI et dirigé réconciliation de données et UAT sur 5+ projets d'entreprise

Stagiaire - SAP Analytics

Blueprint Technologies Pvt Ltd

Juillet 2021 - Décembre 2021, Bengaluru, Karnataka, Inde

- 200+ heures de formation sur SAC, BW, HANA, Python, SQL; contribué aux prototypes et à la documentation ETL/UAT
- Automatisé les 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, Systèmes Distribués, Implémentation SGBD, Deep Learning, Inférence Statistique Avancée

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

- Reproduit expériences Text2KG Bench comparant modèles LLM SOTA pour métriques performance
- Développé pipeline LLM pour génération triplets RDF avec évaluation précision sémantique
- Intégré datasets DBpedia-WebNLG, Wikidata-TekGen, Odeuropa avec ontologies additionnelles

Analyse de Sentiments avec BERT

EURECOM ▪ 2024

- Implémenté modèle BERT pour classification sentiments atteignant 78% F1-score
- Comparé performance avec approches NLP traditionnelles, démontrant résultats supérieurs
- Utilisé Git/GitHub pour contrôle version et gestion projet

COMPÉTENCES

Programmation/AA: Python, SQL, R, scikit-learn, PyTorch, TensorFlow, Pandas, NumPy, SpaCy, HuggingFace, LangChain

Data/BI/Cloud/Devops: Évaluation modèles, ETL, SAP SAC, conception KPI, GCP, AWS, Docker, Git, CI/CD

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