Naveen JOHNSON VALLAVANATT

Nice, France ▪ Email: naveenjohnsonv@gmail.com ▪ Phone: +33 6 02 56 18 12 ▪ LinkedIn: in/naveenjv ▪ Portfolio: naveenjohnsonv.fr

KEY STRENGTHS

Data Scientist with MSc in Data Science (EURECOM) and 2+ years enterprise analytics experience. Expertise in Python, PyTorch, TensorFlow, NLP, deep learning, GenAI, and knowledge engineering. Seeking Data Scientist roles, open to Data Engineer/Analyst roles.

EXPERIENCE

Data Science Research Intern - AI Planning / NLP

ALTEN Labs   March 2025 - August 2025, Valbonne, Alpes-Maritimes, France

- Built LLM pipeline generating formal PDDL tasks executed via Fast Downward planner for playable interactive narratives
- Improved reliability 60% through automated schema/consistency validation and state caching for auditability

Associate Consultant - SAP Analytics

Blueprint Technologies Pvt Ltd   December 2021 - August 2023, Bengaluru, Karnataka, India

- Delivered 20+ SAC dashboards with drill-downs and detailed statistical analyses for pharma and agrochem stakeholders
- Built ETL migrations between SAP ECC and BW/4HANA; designed KPI logic and led UAT across 5+ enterprise projects

Intern - SAP Analytics

Blueprint Technologies Pvt Ltd   July 2021 - December 2021, Bengaluru, Karnataka, India

- Completed 200+ hours across SAC, BW, HANA, Python, SQL; contributed prototypes and ETL/UAT documentation
- Automated data-quality checks using Python/Pandas and SQL extracts

EDUCATION

Master of Science in Computer Science, Data Science Specialization

EURECOM ▪ Biot, France   2023 - 2025

▪ Core subjects: Machine Learning, Reinforcement Learning, Deep Learning, Advanced Statistical Inference, DBMS

Diploma in Data Science

Indian Institute of Technology, Madras ▪ Chennai, Tamil Nadu, India   2021 - 2024

Bachelor of Technology in Electronics and Communication Engineering

National Institute of Technology, Calicut ▪ Kozhikode, Kerala, India   2016 - 2020

PROJECTS

Knowledge Engineering in the LLM Era   EURECOM ▪ 2024

- Reproduced and extended Text2KGBench (Text-to-Knowledge-Graph Benchmark) experiments comparing state-of-the-art LLM models for RDF triple extraction, evaluating precision, recall, and F1-score
- Built ontology-aware LLM pipeline: GPT-4o achieved AUC 0.89 vs Qwen2.5 32B; integrated DBpedia-WebNLG, Wikidata-TekGen, Odeuropa dataset with CIDOC-CRM ontologies to evaluate semantic accuracy/hallucination reduction

Sentiment Analysis using BERT   EURECOM ▪ 2024

- BERT-based tweet classification achieving 78% F1-score; outperformed traditional NLP approaches (TextBlob, VADER) via comprehensive preprocessing

SKILLS

Technical: 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

ML Techniques: Time series (ARIMAX, LASSO), computer vision, RL, optimization, feature engineering

Specialized: NLP, GenAI (rompt engineering, RAG, fine-tuning), time series, computer vision, RL, knowledge graphs

Languages: English - Proficient (C2), French - Intermediate (B1)