Vansh Raja

+91 70219 54565 | vanshraja32@gmail.com | vanshraja.me

ACADEMIC QUALIFICATION

Bachelor of Engineering in Computer Science with specialization in AI and ML, Vellore Institute of Technology, Bhopal, India; Aggregate/CGPA: 8.89/10; Jul'25

INTERNSHIPS

Security Analyst Intern, Virtual Galaxy Infotech Ltd., Nagpur **[Oct '25 – Current]**

Virtual Galaxy Infotech Ltd (VGIL) is an established provider of banking and enterprise software, now enhancing its portfolio with advanced security and AI-supported technologies.

- Built a RAG-based legal drafting system: 5,000+ judgments ingested into PGVector; PDF→metadata→chunking→embeddings pipeline; hybrid retrieval (vector + PostgreSQL FTS); Streamlit interface with petition-aware retrieval and draft generation.
- Built key components of a cross-platform DLP system by developing Windows/Linux agents for directory, clipboard, and PII monitoring, and enabling real-time event forwarding to a FastAPI backend with PostgreSQL, Redis, and OpenSearch for centralized incident logging and policy-based actions.

Data Analyst Intern, PreProd Corp, Bangalore **[Nov '23 – Jan '25]**

*PreProd Corp is an **AI Startup** that empowers individuals and businesses to build sustainable AI through upskilling, consulting, and product innovation.*

- Designed and implemented an end-to-end MLOps pipeline using Kafka, DVC, MLflow, PySpark and Scikit-Learn for a Food Demand Forecasting project. *[Nov '23 – Feb '24]*
- Contributed to PreProdCorp's Opensrc' initiative by developing demo projects using tools like PySpark, RAG, GAN, MobileNet, Redis, and CouchDB. *[Mar '24 – Jul '24]*
- Designed, organized, and executed the 2-day "Buildathon" Hackathon at VIT Bhopal; developed the MVP for the AutoML app demo, mentored participants, and assisted in evaluating final submissions. *[Aug '24]*
- Led 20 interns across 6 teams as **Team Leader**, guiding the development of 30+ demo projects under the PreProdCorp Opensrc' initiative using PySpark, Pandas, Polars, CV, RAG, and GANs. *[Aug '24 – Jan '25]*
  Acted as Scrum Master, conducting sprint meetings, ensuring timely delivery, and promoting MLOps best practices.

Tools Used: PowerBI, Tableau, Apache Kafka, Spark, PySpark, Scikit-Learn, DVC, SQLite, MongoDB, Redis, MLFlow, Flask, Pandas, NumPy, and Pickle.

PROJECTS

Legal Drafting System *[https://github.com/Vansh-Raja/LegalDraftingSystem]* **[Oct '25]**

Built a RAG-based legal assistant with two modes: petition-drafting (upload petition, retrieve case-law, generate draft replies) and legal-research chat over a PGVector-indexed corpus of Indian judgments. Implemented ingestion pipelines for text and metadata, hybrid retrieval (vector + FTS), LLM query planning, and context assembly, deployed through a Streamlit interface with multiple model backends.

Tools Used: Python, PostgreSQL, PGVector, LangChain, Streamlit, Ollama, OpenAI/OpenRouter LLMs, pdfminer.

Sunshine Bot *https://github.com/Vansh-Raja/SunshineBot* **[Sep '25]**

Built an AI voice assistant on a Twilio phone number that identifies patients by phone and manages hospital appointments via Google Calendar + SQLite cache, handling booking/changes/history with doctor-specific, conflict-free slots.

Tools Used: Python, Flask, Google Calendar API, Twilio, ElevenLabs Webhooks, SQLite/Postgres, Waitress.

Food Demand Forecasting *[https://github.com/Vansh-Raja/Food-Demand-Forecasting]* **[Nov '23 – Feb '24]**

Designed a time-series model to forecast food demand with high accuracy. Implemented an MLOps pipeline to enhance project efficiency. Significantly improved food demand prediction accuracy, optimizing inventory management and reducing food wastage.

Tools Used: PowerBI & Tableau, Apache Kafka, Spark & PySpark, Scikit-Learn (SkLearn), DVC (Data Version Control), SQLite, MongoDB, MLFlow, Flask, Pandas, NumPy, PyMongo, Pickle

CERTIFICATIONS

- "Privacy and Security in Online Social Media" by NPTEL, Apr' 24
- "Artificial Intelligence and Machine Learning", by Google Developers and SmartInternz, Nov'23
- "Applied Machine Learning in Python" by University of Michigan, Feb' 23

TECHNICAL SKILLS

ML Frameworks: Python, Scikit-learn, TensorFlow, PyTorch, H2O AutoML, PyCaret
NLP & CV: Transformers, LangChain, GANs, MobileNet, CV Pipelines, TPOT
MLOps & Data: MLflow, Docker, FastAPI, PySpark, Kafka, MongoDB