VIKRAM RAJA

📞 +91 8825615460 📍 Chennai, TN, India 🌐 Portfolio ✉ vvraja2301@gmail.com in linkedin

EDUCATION

VELS INSTITUTE OF SCIENCE & TECHNOLOGY   Jun 2019 - May 2023
Bachelor of Engineering Computer Science and Engineering,   Chennai, TN

EXPERIENCE

FOUNDING ENGINEER & SENIOR ARCHITECT   Mar 2025 - Present
KAMUIT   Dallas, TX

- Designed geospatial ETL pipelines with Python, FastAPI, PostgreSQL/PostGIS, and Uber H3 indexing, powering detour-aware ride matching for 50K+ trips. Implemented spatial joins, polyline overlap, and ETA-based detour scoring, shrinking driver pools 65%, cutting detours 22%, with < 500ms latency. Deployed with Docker/Kubernetes and orchestrated workflows via Airflow + Celery
- Built fault-tolerant streaming infra with multi-passenger batching heuristics, Redis queues, distributed locks, and Stripe Connect APIs, reducing fallback recovery 37% and achieving 99%+ uptime in MVP demos

Full-stack AI Engineer (Python)   Apr 2025 - Present
Intemo Technologies Pvt   Chennai, TN

- Automating complex manual data entry processes in ERP systems through AI-driven document processing, enhancing operational accuracy and efficiency for logistics & freight forwarding clients.
- Developed an end-to-end automation pipeline using OCR, LLMs (ChatGPT & Gemini), and FastAPI that extracts, classifies, and structures emailed documents into JSON, auto-filling ERP fields via Selenium/Playwright with PostgreSQL integration—fully eliminating human input.
- Deployed on AWS using Docker containers, CI/CD, and asynchronous queues to handle multi-tenant loads efficiently and maintain 99.9% uptime.

ML and Model Engineer   Jul 2024 - Apr 2025
Intemo Technologies Pvt   Chennai, TN

- Developed an SVR-based model for predicting flight departure and arrival times (ETD/ETA) using actual data (ATD/ATA), improving scheduling and reducing delays in aviation operations.
- Performed full ML lifecycle — data ingestion, preprocessing, hyperparameter tuning, and deployment on AWS EC2/S3 — with model evaluation using MAE and R^2 metrics.

Full Stack Developer II   Aug 2023 - Jul 2024
Intemo Technologies Pvt   Chennai, TN

- Developed a live container tracking system delivering real-time logistics milestones; built REST APIs using Java Spring Boot, PostgreSQL, Kafka Debezium, and ZooKeeper for event streaming and synchronization.
- Gradually took ownership of core modules and deployments on AWS EC2, containerized with Docker, and independently built 80% of the React.js frontend for real-time tracking dashboards.

PROJECTS

UNDERSTANDING DISEASE HETEROGENEITY(GitHub)   Jan 2025 - May 2025

- Conducted EDA and multivariate analysis on 50K+ synthetic ALS patient records; applied PCA/autoencoders to cut features 120→22 (82% reduction, 93% variance retained) and clustered patients (silhouette > 0.71) into 3 subtypes, deploying results in Plotly dashboards for clinician validation and designing subtype-targeted trials.

IMAGE DIFFUSION FOR SYNTHETIC DATA GENERATION(GitHub)   May 2025 - Oct 2025

- Built a DDPM pipeline with baseline and attention-enhanced U-Net (linear attention, ResNet blocks, group norm, SiLU), on 10K+ Oxford Flowers; used beta scheduling + sinusoidal embeddings to cut reconstruction error 21% and latent size 78% vs. GANs.
- Deployed via Streamlit GPU app for batch sampling (60 images in 4 mins); achieved FID 12.4 (–37%) and IS 4.8..

TECHNICAL SKILLS

Languages & Frameworks: Python, SQL, Java, Springboot, Postgres, Bash, JavaScript, Node.js, React.js; FastAPI, Flask, Django, PyTorch, TensorFlow, Scikit-learn, LangChain, Hugging Face, spaCy.

AI/ML Development: Deep Learning, Transformers, Diffusion Models, GANs, RAG Pipelines, NLP, Time-Series Forecasting, Anomaly Detection, Autoencoders, Clustering (PCA, K-Means), and Model Optimization (ONNX Runtime, Quantization, GPU/CUDA Acceleration).

MLOps & Deployment: AWS (EC2, S3, Lambda, ECR, CloudWatch), Docker, Kubernetes, Airflow, CI/CD (GitHub Actions, GitLab CI), Model Serving via FastAPI + Uvicorn/Gunicorn; API Security with OAuth2, JWT, and Role-Based Access Control (RBAC).

Data Engineering: PostgreSQL, MySQL, MongoDB, Kafka, Debezium, Spark, Airflow, dbt, Snowflake, Redshift, NoSQL, ETL/ELT Pipelines, Vector Databases (FAISS, ChromaDB, ElasticSearch).

Visualization & Analytics: Power BI, Tableau, Plotly, Matplotlib, Seaborn, Dash; strong EDA, feature engineering, and research presentation skills.

Dev Practices: Microservices, Distributed Systems, API Design, Unit Testing, Git, Agile/Scrum, and Cloud-Native CI/CD workflows.