

Maruri Sai Rama Linga Reddy

+91 7893865644 | sairam.maruri@gmail.com | Portfolio | www.linkedin.com/in/sairam-8b31b9275/ | github.com/sairam3824

EDUCATION

Vellore Institute of Technology

Bachelor of Technology in Computer Science and Engineering | GPA: 8.24/10

Coursework: Data Structures & Algorithms, Operating Systems, OOP, Software Engineering (Design Patterns, System Design, Computer Networks, Deep Learning)

Andhra Pradesh, IN

Sep 2022 – Aug 2026

TECHNICAL SKILLS

Programming Languages: Java, Python, C++, SQL

AI/ML: RAG (Retrieval-Augmented Generation), LLMs (OpenAI, Claude, QWEN, BART, BERT, T5), LangChain, MCP Server, Neural Networks, Supervised / Unsupervised Learning, AI Agents, Prompt Engineering

Libraries: PyTorch, TensorFlow, Keras, scikit-learn, Hugging Face Transformers, NumPy, Pandas, Seaborn

Databases: MySQL, MongoDB, Redis, Vector Databases (FAISS)

Cloud/DevOps: AWS (S3, EC2, Lambda, Bedrock), Kubernetes, Docker, Git, Terraform, CI/CD Pipelines

Tools: VS Code, Colab, Jupyter, Thunder Client, Tableau, Power BI

CERTIFICATIONS

AWS Certified: Cloud Practitioner

Oracle: Generative AI Professional

Microsoft: Azure Data Fundamentals

AWS Academy: Cloud Foundations, Architecting

Coursera : Meta - Introduction to Databases, IBM - Machine Learning with Python

PROJECTS

Modular Django AI Platform for Research Paper Workflow and Collaboration

May 2025 – Present

- **Served as AI Engineer, Architect, Tester, and Optimizer**, leading design and deployment of an end-to-end research portal, reducing AI pipeline latency by 30% and boosting system scalability.
- Engineered and deployed production-grade multi-role platform on AWS (orravyn.cloud), supporting 4 roles (Admin, Moderator, Publisher, Reader), 10+ permission-restricted views, and an invite flow validated with 15+ accounts.
- Accelerated search efficiency by 45% through multi-field filtering (title, abstract, author) and 4 sorting modes, enabling 100+ paper queries with real-time chatbot support.
- Architected the **Yggdrasil AI Assistant** using RAG: indexed 100+ PDFs with FAISS and vector database embeddings, enabling retrieval of top 4–5 relevant papers per query.
- Generated multi-document summaries with LLMs and powered future recommendations Lambda via OpenAi. Optimized Claude API chatbot response latency by 25% through LangChain orchestration.

Technologies: Python, BART, FAISS, LangChain, Vector Database, OpenAI GPT, Claude API, AWS, Django, Django REST Framework, Bootstrap, ML, RDS

Bone Fracture Classification

Feb – Mar 2025

- Implemented PyTorch model using WideResNet to classify bone fractures in X-rays, achieving 92% accuracy.
- Applied data augmentation to enhance generalization, reducing overfitting by 20%.
- Visualized performance with Matplotlib and Seaborn, reducing misclassification by 15%.

Technologies: Python, PyTorch, WideResNet, Matplotlib, Seaborn

Customer Churn Prediction

Sep – Nov 2024

- Developed a large-scale system design for a churn prediction model using Random Forest and machine learning techniques, achieving 99% accuracy on 100,000+ records, reducing churn by 25%.
- Reduced dataset noise by 15% through EDA and preprocessing to improve model reliability.
- Visualized insights with Pandas and Seaborn, enhancing customer retention strategies by 30%.

Technologies: Python, Random Forest, scikit-learn, Pandas, Seaborn

EXPERIENCE

Photon Club, VITAP University

Administrator

Andhra Pradesh, IN

Feb 2024 – Apr 2025

- Drove a 15% increase in VTAPP membership by managing marketing strategies for two major events, enhancing the organization's visibility and attracting new members through targeted outreach initiatives.