

CHARAN SAI PONNADA

+91-8639269120 | charansaiponnada06@gmail.com | [LinkedIn](#) | [GitHub](#)

TECHNICAL SKILLS

Programming: Python, Java, C, SQL, JavaScript

AI/ML: PyTorch, TensorFlow, Keras, SciPy, Model Fine-Tuning (LoRA), Evaluation Metrics

LLM Systems: Retrieval-Augmented Generation (RAG), Vector & Non-Vector Retrieval, Prompt Engineering, LLM Evaluation, LangChain

Backend & APIs: FastAPI, Flask, Node.js, REST APIs, API Integration, Microservices

Frontend: Next.js, React (basic integration for AI systems)

Data Engineering: Pandas, NumPy, PySpark, Dask, MySQL, MongoDB

Cloud & Deployment: AWS, GCP, Model Serving, Scalable Inference Systems

Tools & Automation: Git, Docker (basic), UiPath RPA, Workflow Automation

EXPERIENCE

AI Engineer Intern

Jan 2026 – Present

Aynstyn Technologies Pvt. Ltd.

Remote

- Engineered and deployed AI-powered SaaS modules for interview simulation, candidate assessment, and voice evaluation, supporting 3+ production workflows
- Designed and implemented an LLM evaluation pipeline, validating model performance across 200+ structured test cases and improving evaluation coverage by 2x
- Optimized prompt architectures and inference workflows, reducing unstable outputs and improving response consistency across edge cases
- Integrated LLM-based components into scalable SaaS infrastructure, enabling real-time inference and supporting concurrent user interactions

RESEARCH

Vision-Language Assistive Navigation System

2025

Primary Author — ISAECT 2025

DOI

- Built multimodal assistive navigation system using BLIP vision-language model for scene understanding
- Fine-tuned BLIP via 3-stage LoRA training on a 427-scene Indian traffic dataset
- Achieved +15.6% improvement across BLEU, METEOR, ROUGE, and semantic similarity
- Developed full ML pipeline covering dataset design, training workflow, evaluation framework, and model release
- Published:** HuggingFace Model | GitHub Code

PROJECTS

VIVIRITY Intelli-Credit — Credit Risk Intelligence System

GitHub

- Built end-to-end AI credit intelligence system; secured 2nd place at IIT Hyderabad AI/ML Hackathon (YUVAAN 2026)
- Reduced financial analysis time from days to minutes by processing 500+ page annual reports in under 5 minutes
- Developed VectorLess RAG pipeline achieving 98% lower API overhead vs embedding-based retrieval systems
- Extracted 50+ financial risk indicators and designed multi-agent decision pipeline for loan risk assessment

AyurMind — Domain-Specific RAG System

GitHub

- Developed RAG system over 1200+ structured Ayurvedic knowledge entries
- Achieved 84% Top-3 retrieval accuracy through optimized retrieval pipeline
- Reduced hallucinated responses by 45% using validation layers and filtering mechanisms
- Improved expert-rated answer quality from 3.1 to 4.4 / 5

EDUCATION

Velagapudi Ramakrishna Siddhartha Engineering College

B.Tech in Artificial Intelligence and Data Science — CGPA: 8.76

Vijayawada, India

2023 – 2027