

# Saikiran Bondi

B.Tech

CSE with Major in AI

Indian Institute Of Information Technology,

Design and Manufacturing, Kancheepuram

## EDUCATION

- 
- |  |            |
|--|------------|
| •Indian Institute Of Information Technology, Design and Manufacturing, Kancheepuram        | 2025       |
| B.Tech Computer Science and Engineering with Major in Artificial Intelligence(Distinction) | CGPA: 9.12 |
| •SASI Junior College   | 2021       |
| Board of Intermediate Education, Andhra Pradesh  | CGPA: 9.78 |
| •Holy Angels E.M School  | 2019       |
| Board of Secondary Education, Andhra Pradesh   | CGPA: 10   |

## EXPERIENCES

- 
- |  |                             |
|--|-----------------------------|
| •Ruckus Networks   | Chennai, India              |
| Software Engineer (Full Time).   | July 2025 – Present         |
| – Started as a Full-Time Software Engineer in the Infrastructure Team.   |                             |
| • GoLLM  | Remote                      |
| Full Stack and AI Intern.  | December 2024 - March 2025  |
| – Built a Video-RAG system for timestamp-based retrieval in videos using Whisper, Azure Speech-to-Text, Qdrant, and Gemini AI.   |                             |
| – Developed LLM-based document automation pipelines for AI-driven reporting.   |                             |
| – Designed and deployed REST APIs for scalable backend operations and seamless data processing.  |                             |
| •ESAB India  | Chennai, India              |
| Full Stack and AI Intern.  | June 2024 - August 2024     |
| – Implemented an LLM-powered retrieval system for 25,000+ industrial documents using multi-modal RAG.  |                             |
| – Worked with embedded C and hardware interfaces to collect real-time data from industrial machines.   |                             |
| – Contributed to a project on material lifetime prediction using various ML/DL models and deployed TFLite models on edge devices.  |                             |
| •IIITDM Kancheepuram   | Chennai, India              |
| Research Intern.   | August 2024 - November 2024 |
| – Co-authored a paper on continual learning for chest X-ray disease classification titled " <i>Evaluation of Class Incremental Replay Methods for Lung Diseases Classification</i> ", submitted to CVIP (Paper ID: 444). |                             |
| – Implemented GAN architectures (CycleGAN, Pix2Pix, WGAN) for image synthesis and domain adaptation using PyTorch and related libraries.   |                             |
| – As a continuation, developed a novel hybrid continual learning strategy named EWC-Replay.  |                             |

## PROJECTS

- 
- |   |
|---|
| •GoLLM Survey Analytics Tool  |
| Developed an AI-powered survey analytics platform with LLM-based insights, sentiment analysis, and automated reporting.       |
| – <b>Tech Stack:</b> FastAPI, React, PostgreSQL, Redis, LangChain.  |
| – Built LLM-powered analytics workflows, clustering, and sentiment classification models.                                     |
| – Developed REST APIs using FastAPI for scalable data processing and backend integration.                                     |
| – Optimized PostgreSQL and Redis caching for efficient retrieval and storage.   |
| •Video-RAG System   |
| Built an AI-driven video search engine for precise timestamp retrieval from transcribed video datasets.                       |
| – <b>Tech Stack:</b> Azure Speech-to-Text, Whisper, Qdrant, Gemini AI, Redis, FastAPI.  |
| – Developed real-time transcript indexing and semantic search across multiple video sources.                                  |
| – Built REST APIs for query-based video timestamp retrieval using Qdrant and LLMs.  |
| – Optimized backend performance using Redis caching for low-latency search.   |
| •MCP Reddit Agent – Generalizable Automation Bot  |
| Built a modular Reddit automation system using MCP and Docker, capable of scaling to other platforms.                         |
| – <b>Tech Stack:</b> MCP, PRAW, LangChain, Celery, PostgreSQL, FastAPI, Google A2A, 11 Labs, Docker.                          |
| – Developed modular command pipelines to scrape Reddit posts, summarize content using LLMs, and generate automated responses. |
| – Designed for multi-platform generalization with plug-and-play modules (e.g., LinkedIn, GitHub Jobs).                        |
| – Enabled voice output using 11 Labs; fully dockerized for deployment.  |

•**LangGraph Tabular Data Handler**

*Engineered an LLM-powered interface to query structured enterprise data with multi-source alignment and summarization.*

- **Tech Stack:** LangGraph, Pandas, OpenAI API, FastAPI, PostgreSQL.
- Parsed natural language queries into actionable LangGraph workflows that route to relevant structured sources.
- Converted abstract queries into pandas filters, retrieved data, and generated summaries using LLMs.
- Built and deployed multiple LangGraph workflows for diverse tabular tasks.

•**ML Studio**

*Developed an automated ML model selection tool to determine the best algorithm for a given dataset.*

- **Tech Stack:** TensorFlow, Streamlit, Scikit-learn.
- Designed a user-friendly interface allowing users to upload datasets and get model recommendations.
- Automated feature selection and model evaluation pipelines using Scikit-learn.
- Won first place at Vashisht ML Hackathon 2024.

**TECHNICAL SKILLS AND INTERESTS**

---

**Languages:** Python, C, C++, JavaScript, TypeScript, HTML5, CSS

**Backend Technologies:** FastAPI, Django, Node.js, REST APIs

**AI/ML Frameworks:** TensorFlow, PyTorch, LangChain, Langgraph, Google ADK-A2A, Hugging Face, OpenAI, Gemini AI, Crew AI

**Databases:** PostgreSQL, MySQL, MongoDB, Redis, Qdrant

**Cloud Services:** AWS (EC2), Azure (Speech-to-Text, OpenAI)

**DevOps CI/CD:** Docker, GitHub Actions

**Frontend Technologies:** React.js, Next.js, Tailwind CSS

**Developer Tools:** Git, GitHub, Jupyter Notebook, VS Code, Postman

**Soft Skills:** Cross-Functional Collaboration, Problem-Solving, Adaptability, Communication, Analytical Thinking

**Areas of Interest:** LLMs, NLP, RAG Systems, Continual Learning, AI Model Deployment, Backend Performance Optimization

**POSITIONS OF RESPONSIBILITY**

---

•**Joint Core, CS Club - AI Wing - Part of Teaching and Competitive Team** *Aug 2022 - May 2024*

•**Powerlifting Captin, IIITDM Institute’s Powerlifting - Men** *Jul 2023 - Present*

**ACHIEVEMENTS**

---

•**Won Vashisht ML Hackathon 2024:** Led a team of 4,Project Submission : ML Studio Application

•**Won various Coding Contests (Vashisht Tech Fest):** 1st in Prompt Engineering 2025, 2nd in Cryptic Messages 2024