Ananya Naga Raj

Syracuse, NY | annagara@syr.edu | +1 (646) 296-1005 | LinkedIn | GitHub | Portfolio

SUMMARY

Software Engineer & AI/ML Engineer with experience building production-ready LLM-powered applications, RAG architectures, and full-stack systems using Python, React, and cloud technologies. Strong background in machine learning, NLP, and deep learning with hands-on experience in Whisper, Transformers, PyTorch, TensorFlow, vector databases, and REST APIs. Proven ability to deliver scalable, high-impact solutions across AI research and real-world applications.

EXPERIENCE

Software Developer — **Nov 2025 – Present**
iConsult Project, Syracuse University — Syracuse, NY

- Designed and implemented LLM-powered agent workflows using Python, improving system efficiency by 40%
- Applied prompt engineering and RAG techniques to enhance recommendation accuracy by 25%
- Built scalable REST APIs and integrated vector databases to support semantic search and retrieval
- Collaborated with cross-functional teams to deploy production-ready AI features

Graduate Research Assistant — **Oct 2025 – Present**
Biomedical & Chemical Engineering, Syracuse University — Syracuse, NY

- Developed AI-driven data processing and modeling pipelines using Python and TensorFlow
- Optimized large-scale data workflows, reducing processing time by 40%
- Implemented automated data validation and quality assurance frameworks for research datasets

Software Developer — **Mar 2023 – Jun 2024**
Selfy Page Developers Ltd. — Karnataka, India

- Built LLM-based recommendation systems using semantic search and vector embeddings, serving 10,000+ users
- Developed text-to-SQL pipelines enabling natural language querying of relational databases
- Implemented CI/CD workflows and monitoring for AI systems using GitHub Actions

PROJECTS

SafeVoice AI – Harassment Detection System — GitHub | Live Demo

- Built production RAG system with Whisper + NLP achieving 85%+ accuracy for real-time harassment detection, serving thousands of users
- Engineered audio pipeline processing 500+ hours with emotion classification using Transformers and LangChain for context-aware threat analysis
- Implemented Pinecone vector database for semantic search enabling 60% faster incident retrieval with automated response generation
- Developed React frontend with Flask REST APIs integrating real-time WebSocket communication for live threat monitoring dashboard

IoT Attack Detection & Visualization — GitHub | Demo

- Achieved 99.90% accuracy on CICIDS2017 dataset building CNN/LSTM ensemble for IoT attack classification across 12 attack types
- Engineered automated feature extraction pipeline processing 2M+ network records, reducing detection latency by 45% with parallel processing
- Developed interactive dashboard with SHAP explainability and real-time visualization, improving security team response time by 40%

Medical NER with BioBERT — GitHub | Demo

- Fine-tuned BioBERT achieving 92% F1-score for disease and chemical extraction from 15,000+ medical abstracts
- Exposed NER pipeline via REST APIs enabling clinical text processing, reducing manual annotation time by 75%

EDUCATION

Master of Science in Computer Science — **Aug 2024 – May 2026**
Syracuse University, Syracuse, NY — GPA: 3.5/4.0

Bachelor of Technology in Computer Science and Engineering — **Aug 2019 – Jun 2023**
Adichunchanagiri Institute of Technology, India — GPA: 3.48/4.0

TECHNICAL SKILLS

Programming Languages: Python, JavaScript, TypeScript, SQL
Frameworks & Libraries: React, Flask, FastAPI, LangChain, Hugging Face, PyTorch, TensorFlow, scikit-learn
Machine Learning & AI: Machine Learning, Deep Learning, NLP, LLMs, RAG Architectures, Prompt Engineering
Data & Databases: PostgreSQL, MySQL, MongoDB, Redis, Vector Databases (FAISS, Pinecone)
Cloud & DevOps: AWS, Docker, CI/CD, GitHub Actions
Tools: Git, GitHub, Streamlit, SHAP