SRIKANTH REDDY NANDIREDDY

srikanthreddynandireddy@gmail.com | linkedin.com/in/srikanth-reddy-nandireddy
github.com/GodVilan | srikanthreddynandireddy.me | Warrensburg, MO (Relocating)

PROFESSIONAL SUMMARY

AI Engineer with experience in developing and deploying large-language-model applications. Delivered production-grade ReAct agents and retrieval-augmented generation pipelines on AWS, and built end-to-end MLOps workflows using FastAPI, Docker, and MLflow. Certified in Anthropic's Model Context Protocol and proficient in PyTorch, HuggingFace Transformers, and FAISS. Seeking to apply generative-AI expertise to accelerate product innovation and drive measurable business impact.

CORE SKILLS

Generative AI & LLM: ReAct Agents, RAG, Prompt Engineering, FAISS, BGE, BM25, Vector Databases, Sentence Transformers, Google Gemini API, MCP

Deep Learning & NLP: PyTorch, TensorFlow, Keras, Transformers, BERT, DistilBERT, Wav2Vec2, ViT, HuggingFace, CNNs, TF-IDF, Word Embeddings

Machine Learning: Scikit-learn, XGBoost, Random Forest, Logistic Regression, K-Means, Feature Engineering, Model Evaluation

MLOps & Cloud: MLflow, FastAPI, Docker, GitHub Actions, CI/CD, REST APIs, AWS (EC2, S3, Athena, Glue, SageMaker)

Data Analysis: Pandas, NumPy, SciPy, Statsmodels, EDA, Time Series Forecasting, Hypothesis Testing

Programming: Python, SQL (MySQL, PostgreSQL)

Visualization: Matplotlib, Seaborn, Plotly, Dash, Streamlit

EXPERIENCE

Graduate Laboratory Assistant – Data Science | University of Central Missouri | Warrensburg, MO *Feb 2025 – May 2026*

- Designed and delivered 15+ ML laboratory sessions per semester for 40+ students, covering Python-based supervised learning, scikit-learn model pipelines, and end-to-end evaluation workflows from data ingestion through prediction
- Accelerated student proficiency across EDA and visualization workflows using Pandas, NumPy, and Matplotlib, while maintaining a 4.0 GPA concurrently across full-time graduate coursework and research responsibilities

Data Analytics Virtual Internship | AICTE EduSkills – Supported by AWS Academy | Remote *May 2023 – Jul 2023*

- Engineered an end-to-end cloud analytics pipeline integrating Amazon S3 for storage, AWS Glue for ETL automation, and Athena for serverless SQL querying, enabling automated metadata management across 10K+ semi-structured records
- Built predictive models using Python and Amazon SageMaker to identify actionable dataset trends, completing a structured 10-week data analytics program supported by AWS Academy curriculum

Data Scientist Intern | Swecha | Hyderabad, India *May 2022 – Jun 2022*

- Improved NLP model reproducibility by 15% across all deployment environments by containerizing Telugu speech recognition pipelines with Docker, eliminating environment-specific runtime failures during production handoff
- Fine-tuned deep learning speech-to-text models using PyTorch, TensorFlow, and NLTK on dialectally diverse Telugu audio data spanning 5+ regional language variants, enabling voice command recognition for an underserved language community

CORE PROJECTS

arXiv Agent – ReAct Research Assistant | GitHub

Python · FAISS · BGE · BM25 · Gemini 2.5 · Streamlit

- Achieved MRR@5 = 0.990 and Context Precision = 1.000 on 100 curated QA pairs by engineering a ReAct agent with multi-hop query decomposition, iterative retrieval, and a self-critique loop that refines weak answers before delivery
- Outperformed BM25 baseline by 7.3× in answer relevance (0.912 vs. 0.125) by implementing BGE-large-en (1024-dim) embeddings with FAISS IndexFlatIP, benchmarked against MiniLM and MPNet across 3 chunk sizes and 4 retrieval depths
- Delivered a production-ready research assistant with 5 agent tools (semantic search, BM25 fallback, live arXiv fetch, summarize, compare), conversation memory, and persistent research notes via Streamlit UI

Real-Time Credit Card Fraud Detection System | GitHub

Python · Scikit-learn · XGBoost · MLflow · FastAPI · Docker · AWS (EC2, S3) · GitHub Actions · CI/CD

- Deployed a Dockerized FastAPI service with REST endpoints on Amazon EC2 for real-time fraud prediction and batch scoring, fully automated through a GitHub Actions CI/CD pipeline
- Versioned and compared 3 fraud classification experiments (Logistic Regression, Random Forest, XGBoost) in MLflow Model Registry with all artifacts stored in Amazon S3, enabling complete experiment reproducibility

- Addressed class imbalance in fraud classification using SMOTE oversampling with ROC-AUC as the primary optimization metric across realistic imbalanced transaction datasets

Multimodal Emotion Recognition using Cross-Modal Transformers | GitHub
Python · PyTorch · HuggingFace Transformers · BERT · Wav2Vec2 · ViT

- Attained 59.96% accuracy and 60.66% weighted F1-score on 7-class emotion recognition by designing a Transformer fusion encoder (4 layers, 8 attention heads) that aligns BERT, Wav2Vec2, and ViT representations across text, audio, and visual modalities
- Resolved 7-class label imbalance on the MELD dataset by applying class-weighted cross-entropy loss, improving minority emotion class recognition relative to standard training baselines

US COVID-19 Mortality Forecasting System | GitHub
Python · Pandas · NumPy · Statsmodels · XGBoost · Prophet · Plotly Dash · Docker

- Outperformed Prophet and XGBoost baselines in daily mortality forecasting by implementing Seasonal SARIMAX (1,1,1)(1,1,1,7) on 3,300+ records, achieving holdout MAE of 138.14 and cross-validation MAE of 137.86
- Constructed 6 time-series features — lag variables, rolling statistics, and calendar features — capturing seasonal mortality patterns that reduced overall forecasting error relative to non-engineered baselines
- Deployed a Plotly Dash dashboard with 30-day mortality forecasts and 95% confidence intervals using Docker on Render, enabling interactive public health trend visualization

EDUCATION

M.S. – Data Science & Artificial Intelligence| GPA: 4.0 / 4.0 *Aug 2024 – May 2026*
University of Central Missouri | Warrensburg, MO
Coursework: *Machine Learning, Artificial Intelligence, Neural Networks & Deep Learning, NLP, Statistical Foundations of DS & AI, Advanced Data Visualization*

B.S. – Computer Science & Engineering (Honours) | GPA: 9.06 / 10.0 | First Class with Distinction *Nov 2020 – Apr 2024*
CMR College of Engineering & Technology | Hyderabad, India

CERTIFICATIONS

- Introduction to Model Context Protocol | Anthropic | Apr 2026
- AWS Academy Graduate – Data Analytics & Cloud Foundations | Amazon Web Services | Jun 2023
- Foundations: Data, Data, Everywhere (Google Data Analytics Certificate) | Google / Coursera | Feb 2023
- Artificial Intelligence & Machine Learning | Blackbuck Engineers Pvt. Ltd. | Dec 2021

HONORS & AWARDS

- Graduate Student Achievement Award – UCM
- Finalist — IICC National Coding Competition (Top 1% of 100,000+ participants)
- 5★ HackerRank | 4★ CodeChef in Data Structures & Algorithms
- GeeksforGeeks POTD Streak — 255 consecutive days solving algorithmic problems