Venkata Sai Revanth Damisetty

revanth.damisetty@gmail.com | **Cornell University** | (607) 339-6540 | New York, USA
linkedin.com/in/revanth-damisetty | leetcode.com/u/revanth_damisetty | github.com/revanth-damisetty
Open to Summer 2026 oppurtunities in ML, AI Engineering, and Data Science.

Education

Cornell University — Ithaca, NY
Masters in Applied Statistics (Data Science) — Aug 2025 – May 2026
Amrita Vishwa Vidyapeetham — Coimbatore, India
Bachelors – Computer & Communication Engineering | **Department Rank: III** — Aug 2020 – May 2024

Technical Skills

Software Tools: Python, R, Git, Flask, Docker, Angular, Streamlit, FastAPI, Tableau
Methodologies: Experiment Design (A/B Testing), Feature Engineering, Model Inference & Finetuning, Visualization
Academic Exposure: Statistical Modeling, Experiment Design, Prompt Engineering, Agentic AI, LLM Evaluation, Observability
AI & Data Science Frameworks: Numpy, Pandas, Scikit-learn, TensorFlow, Keras, PyTorch, Hugging Face, LangChain, LlamaIndex
Databases & Cloud Computing: SQL, MongoDB, ChromaDB, Spark, Hadoop, CUDA, AWS (EC2, SageMaker, Bedrock)
LeetCode: Solved 150+ problems | Rating: 1696 (**Top 15%**)

Certifications

- **AWS ML Engineer – Associate**
- **Machine Learning Specialization** – Stanford, Coursera
- **Generative AI for Data Scientists** – IBM, Coursera

Work Experience

Software Development Engineer, GE Healthcare — Aug 2024 – Jun 2025
– Built an **Generative AI system** to assist sales teams in hospital compliance questionnaires, reducing manual submission times by **40%** (KPI).
– Processed **15,000+** unstructured Word and Excel documents using a **Python-based parsing automation pipeline** with semantic pattern strategies like **TF-IDF embeddings** and **cosine similarity** to embed into a **ChromaDB VectorStore**.
– Implemented a **Retrieval-Augmented Generation (RAG)** pipeline using **LlamaIndex** retrieval, **Ollama** Large Language Model (LLM) and **Langchain** Orchestration for context-based responses within enterprise documentation and good Retrieval Quality.
– Deployed a **secure, Dockerized** service, exposing ranked inference via **REST APIs** into a **Microsoft Office plugin** (forms – dropdown options).

Intern, GE Healthcare — Jan 2024 – Jul 2024
– Built a **Dependency Management System** to diagnose **anomalous deployment failures** caused by low-level dependency mismatches across C++, Java, and Python introduced by feature updates that are not captured by **SBOM**, effectively cutting **release debugging time to 75%**.
– Designed a **hybrid persistence layer** with **MongoDB** dependency trees and **SQL** version metadata extracted by a **Bash** system-level crawler.
– Implemented **dependency graph diffing** between releases to surface new paths and high-impact dependency shifts for failure investigation.
– Designed an **AI-assisted root cause analysis** workflow that reasons over dependency diffs and contextual metadata to generate **ranked hypotheses**, evidence-backed explanations, and actionable validation steps for debugging in a CI/CD environment.
– Delivered a **Dockerized** full-stack platform (**Angular** UI + **Flask** REST APIs), provided demos to **3 modality teams**, incorporated feedback.

Projects

SignStream: Real-Time American Sign Language (ASL) Recognition (Deep Learning + Computer Vision)
– Orchestrated an end-to-end deep ML system for ASL gesture recognition, achieving **99.16% classification accuracy** across 29 classes.
– Transformed gesture images into **3D hand landmarks** using **Mediapipe** library along with **tilt-altered dataset augmentation**, to improve robustness under varied camera angles and hand orientations.
– Trained a multi-layered deep neural network on **PyTorch** and evaluated reliability using **confusion matrices, per-class recall and ROC analysis** to ensure consistent performance across correlated ASL Signs.
– Performed model deployment into a **low-latency containerized platform (<300 ms per frame)** on **Streamlit** using **FastAPI** inference calls, **OpenCV-based** asynchronous frame capture and prediction throttling for real-time performance.

Systemic Risk Analytics: U.S. Financial Crisis (2007–2009)
– Investigated systemic risk propagation across housing, equity, credit, and labor markets, identifying **housing price declines as 6 –12 month leading indicators** of equity drawdowns and mortgage delinquency spikes to support early-warning risk assessment.
– Designed and implemented an end-to-end **feature engineering pipeline** using **Numpy, Pandas**, integrating macro-financial data (2003 –2011) from **FRED** and **Yahoo Finance** using **temporal alignment, resampling, and interpolation** to construct an analytical panel.
– Conducted Causal Analysis using **lagged regression models, rolling-window correlation analysis, and year-over-year differencing**, translating into dashboard-ready metrics and stakeholder-facing visualization in **Tableau**.

Publications

- **Optimizing Machine Learning Pipelines: Scalable Mobile-Based Tamil Sign Language Recognition** – IEEE NKCon 2024
- **Smart Detection of Cyber Threats: A ML Approach to Real-Time Network Traffic Analysis** – IJSREM 2024