

# Sriman Donthireddi

(317) 514-0162 | srimanvd@gmail.com | linkedin.com/in/srimand | github.com/srimanvd | U.S. Citizen

## EDUCATION

### Purdue University

Expected May 2026

B.S. in Computer Science **GPA: 3.85/4.0**

- **Honors:** Scholars List (2022–2024), 21st Century Scholar
- **Coursework:** Data Structures & Algorithms, Software Engineering, AI, Deep Learning, Operating Systems, Databases, Data Engineering, Computer Architecture, Theory of Computation, Linear Algebra, Discrete Math

## EXPERIENCE

### Co-founder

Aug 2025 – Present

ComputeAI (*computeai.app*) | Indianapolis, IN

- Co-founded **ComputeAI**, a data centric AI startup that turns noisy Common Crawl STEM content into pretraining, SFT, and RL datasets for LLMs and collects grounded desktop interaction data with chain of thought to train agentic models.
- Built the first full data pipelines for these two tasks, delivering training ready datasets by designing dataset formats, implementing Common Crawl parsing and desktop event logging with millisecond level timestamps and aligned chain of thought, and adding automated checks to remove low quality samples.

### Computer Vision Researcher

Aug 2025 – Present

The Data Mine & V2X | Indianapolis, IN

- Developing an image recognition system on a **100k+ image dataset** to identify vehicle license plates in varied conditions.
- Designing a pipeline to process **3,000+ images/hour** using **YOLO, Vision Transformers, CNNs**, and **PyTorch**, with deployment planned for real-world **automated tolling and enforcement**.

### Software Engineer Intern I & II

May 2024 – Nov 2025

Erie Insurance Group | Erie, PA

- Rebuilt a legacy life-insurance quoting system on a PERN stack (Postgres, Express, React, Node), used by **6,000+ employees/agents**, and containerized/deployed it with **Docker** and AWS (**ECS/Fargate, Lambda, Amplify**) to improve scalability and reliability.
- Developed new life-insurance product features and refactored legacy code (metadata, dashboard deprecation) in an Agile team, improving usability, maintainability, and data accuracy.

### AI/ML Research Lead

Aug 2024 – May 2025

The Data Mine & Corteva Agriscience | Indianapolis, IN

- Led a team of 8 to build an AI chatbot that let scientists instantly search and analyze **5,000+ product SDS/Label PDFs**.
- Developed an **innovative graph-based Retrieval-Augmented Generation (RAG) system**, enabling large-scale Q&A across thousands of documents beyond traditional context window limits.
- Delivered estimated **multi-million dollar annual savings** in labor and hiring costs by cutting research time **95%**, transforming weeks of manual review into minutes.

## PROJECTS

### Phantasia — Best Overall Winner, Purdue INDY Hackathon 2025

Mar 2025

- Led a team of 4 to build an **educational platform** with PDFs, LaTeX editing, graphing, and AI-powered **Manim** video generation plus screenshot-based **AI Q&A** via a multi-agent framework (**MCP + OpenAI, Gemini, DeepSeek**), cutting STEM content creation and study time from days to minutes.

### Customer Order — Cloud Data Engineering

Feb 2025

- Built **ETL/ELT** pipelines using **Azure Data Factory, Cosmos DB, PostgreSQL**, and **Blob Storage** to support both OLAP and OLTP workloads, reducing ingestion time by **70%** through optimized data flows and CI/CD for a fault-tolerant, scalable architecture.

## SKILLS

**Languages:** Python, Golang, Rust, C, C++, JavaScript/TypeScript, Java, SQL, HTML, CSS

**Frameworks/Databases/Cloud & Tools:** React, Svelte, Node.js, Next.js, Electron.js, Tauri, PostgreSQL, MongoDB, MySQL, Pinecone, PySpark, Docker, Git/GitHub, Linux, AWS (ECS, Lambda, SageMaker), Azure (ADF, Fabric), Terraform, Kubernetes, Apache Airflow, PowerBI, CI/CD

**Machine Learning:** PyTorch, TensorFlow, Hugging Face, LangChain, LlamaIndex, RAG architectures, Adversarial Training, Model Context Protocol, Agent2Agent