Sai Teja **Pusuluri**

Generative AI Expert

158 Brets Ln, Lewis Center, Ohio, 43035

(+1) 740-818-6309 | sai19872000@gmail.com | sai19872000 | Sai Teja Pusuluri

Overview

PhD in Physics with over 8 years of industry experience in Generative AI, Agentic AI Workflows, Traditional ML/AI and MLOps. Proven track record of leading cross-functional teams, driving innovation, and delivering scalable AI solutions. Seeking suitable roles to leverage technical expertise and leadership skills to transform organizations.

Education

Ph.D. Physics (Neural networks)

Aug. 2011 - Feb. 2017

Department of Physics and Astronomy, Ohio University, Athens, Ohio

- GPA: **3.86/4**

Skills

Agentic AI	Agentic Tools, MCP servers, LangGraph, CrewAI, Transformer models, Gemini, Claude, OpenAI, Open Source Models (Llama, Mistral), RAG, Chain-of-Thought prompt engineering.
MLOps	Docker, AWS SageMaker, Amazon ECR, CI-CD, APIs, VLLMs, Python, TensorFlow, PyTorch, Keras, SQL, Bash.
Leadership	Team Management, Mentoring, Cross-functional Collaboration, Stakeholder Engagement.

Work Experience

Manager/Lead - Generative AI

Apr. 2022 - Present

Discover

- Pioneered **agentic AI workflows**, automating customer experience and operational processes and generating comprehensive model documentation via LLM pipelines, resulting in a **75% reduction** in manual effort and improved consistency across deliverables.
- Fine-tuned LLMs for company-specific datasets: applied **LoRA, QLoRA, and quantization** techniques to enhance **local inference throughput by 40% and reduce GPU memory usage by 25%**.
- Implemented **VLLM-based inference optimization**, boosting custom LLM throughput by 50% and slashing latency for production endpoints.
- Mentor and Lead a team of **5 direct and 10 indirect reports**, cultivating talent and ensuring delivery of high-impact AI solutions.
- Designed and maintained robust **MLOps pipelines** with CI-CD, monitoring, and custom API endpoints, ensuring scalable and reliable model operations.

Senior Applied AI - ML Associate

Apr. 2017 - Apr. 2022

J P Morgan Chase

- Developed **fraud detection AI/ML models**, significantly enhancing the bank's ability to identify and mitigate fraudulent activities.
- Engineered **advanced feature extraction** and seasonality methods, refining predictive capabilities and improving model reliability.
- Researched and implemented **scalable neural network and NLP solutions** (CNNs, LSTMs, BERT), advancing the processing of both structured and unstructured financial data.
- Compiled comprehensive technical documentation for end-to-end model development pipelines, ensuring transparency, **audit readiness, and regulatory compliance**.
- Collaborated with business and IT stakeholders to translate requirements into production-ready solutions, deploying and troubleshooting models for **seamless integration**.

Adjunct Professor

Nov. 2023 - Present

Ohio University

- Researched and taught **advanced computer vision and segmentation methods** for organoids, **contributing to cutting-edge academic publications**.
- Developed **agentic AI research workflows**, automating dataset curation and model retraining for organoid segmentation studies.
- Deployed reproducible research environments using **Docker, AWS SageMaker, and ECR**, enabling seamless collaboration and scalability.
- Developed a **Agentic Object Detection Backup Workflow** using a scene-explorer agent, grounding agent, and canvas agent into a tri-agent fallback system, boosting YOLO like detection recall and robustness in complex scenarios.
- **Mentored PhD candidates**, guiding their research on deep learning applications in biology.

Data Scientist

Jun. 2016 - Apr. 2017

Nationwide Children's Hospital

- Developed **deep learning bioinformatics tools** for real-time analysis of infant sensor data, enabling early detection of physiological anomalies.
- Designed and implemented waveform analysis algorithms to identify **feeding disorders** in neonatal patients using **ML-driven pattern recognition**.
- **Applied fast Fourier transform and statistical methods** to extract actionable clinical insights from biosensor signals.
- Collaborated with **pediatric research teams** to translate analytical findings into **improved care protocols and treatment strategies**.
- Documented analytical workflows and presented results in internal reviews, fostering adoption of **AI-driven monitoring solutions.**

AI Citation Builder

Jan. 2025 - Present

Personal Project

- Engineered and implemented an **agentic AI flow** with UI that ingests manuscript sections, autonomously searches academic databases and the web, and generates formatted citations, streamlining literature reviews.
- Implemented deep research capabilities, enabling comprehensive **topic exploration and reference discovery**.

Smart Latex Editor

Jan. 2025 - Present

Personal Project

- Engineered and implemented an **agentic AI flow** with UI that take in custom instruction from user and make edits to LaTex files
- Implemented smart edit and generate capabilities. The app can also **convert regular equations into a LaTex format**.

Publications & Conferences

- Pusuluri et.al., Studying Brain Organoids Survival rates using automated segmentaion methods, 2025 (to be submitted).
- Pusuluri et.al., Image segmentation methods for automated morphological analysis of organoids, 2025 (to be submitted).
- Pusuluri et.al., Dynamics of driven transitions between minima of complex energy landscapes, 2025 (to be submitted).
- Pusuluri et.al., Electrophysiological maturation of cerebral organoids correlates with dynamic morphological and cellular development, **Stem cell reports** 15 (4), 855-868, 2020
- Pusuluri et.al., Cellular reprogramming dynamics follow a simple one-dimensional reaction coordinate, **Physical Biology** 10.1088/1478-3975/aa90e0, 2017
- Pusuluri et.al., Controlling energy landscapes with correlations between minima, arXiv:1611.06127, 2016.
- Pusuluri et.al., Role of deoxy group on the high concentration of graphene in surfactant/water media, 2012, Royal Society of Chemistry (**RSC Advances**).
- **American Physical Society (APS) talk**, Spin glass model for dynamics of cell reprogramming, J47.002, 2015.
- **American Physical Society (APS) talk**, Effects of correlation on complex landscapes, B41.00006, 2016.
- **American Physical Society (APS) poster**, Dynamics driven transitions between minima on landscapes, M1.00278, 2016.
- **American Physical Society (APS) poster**, Controllability of energy landscapes by varying correlations between minima, 2017.