Sanika Chavan

+1 602 516 5693 • sanikac10@gmail.com • LinkedIn • GitHub

EDUCATION

M.S. Computer Science *May 2025*

Arizona State University, Tempe, AZ

PROFESSIONAL EXPERIENCE

Research Assistant, Arizona State University *June 2025*

- Implemented distributed training of large language models (**Qwen 2.5 Math 7B**) with CUDA, with **distributed training** with **A100** GPUs, achieving a 2.8x training throughput increase through system optimization, **parallel processing**.
- Replicated and enhanced **Tiny Zero** reasoning framework using iterative **SFT** instead of **GRPO**, incorporating iterative fine-tuning with correct/incorrect labels (Reinforcement Learning) to improve model performance by 15%.
- Designed a comprehensive evaluation framework with automated metrics tracking, reducing assessment time by 70% while increasing result reliability.
- Engineered SLURM scripts for efficient **resource allocation** and **batch processing**, enabling 24/7 training utilization across shared compute infrastructure.

Software Developer Intern | *Vidyalankar Institute of Technology, Mumbai* *August 2021*

- Designed and implemented a Student Repository System using Java and MySQL with distributed architecture patterns, enabling efficient management of 10+ extracurricular activities across multiple department servers.
- Deployed a scalable system infrastructure with basic load balancing capabilities, facilitating institution-wide adoption within 3 months and supporting concurrent access from various administrative departments.
- Optimized database queries and implemented basic caching strategies, resulting in 2x reduction in administrative overhead and improved response times for student participation monitoring.

TECHNICAL PROJECTS

Multi-Stage Reasoning Framework for VLLMs (Strong Compute Hackathon Winner) *April 2025*

- Engineered a synthetic data generation pipeline for the ARC-AGI-2 challenge benchmark, demonstrating superior general pattern recognition capabilities in our vision-LLM model.
- Structured model responses using three custom token blocks (visual explanation, logical planning, transformation code), which improved interpretability and enabled systematic reasoning on abstract visual patterns.
- Achieved 75+% resolution rate on the training set by optimizing the architecture for LIMO-based reasoning, significantly enhancing the model's ability to generalize across unseen pattern recognition tasks.

Data Leakage in LLMs *December 2024*

- Designed and implemented a detection pipeline to identify key data leakage scenarios in large language models.
- Evaluated five diverse models (e.g., GPT-4o, Claude-3.5, Llama-3.2) to analyze susceptibility to leakage, revealing key insights about model generalization, reasoning, and reliance on training data patterns.
- Developed and validated the Relevant Information module, leveraging techniques like Part of Speech tagging and entailment checks to detect data leakage and highlight the need for robust training processes and dynamic evaluation benchmarks.

KAN You See the Bias? A visualization of Comparative Study of MLPs vs. KANs *December 2024*

- Faced the challenge of bias in machine learning models impacting human-centric decisions. Built a **storytelling visualization** using ADULT and WESAD datasets, integrating **heatmaps** and **bias distribution plots** to communicate these effects. Improved user engagement by 20%, effectively illustrating the real-world consequences of biased models.
- Analysed and compared Multi-Layer Perceptron (MLP) and Knowledge-Aware Network (KAN) architectures, demonstrating a 10% improvement in generalization and reduced bias with KANs. Visualized performance and fairness metrics across sensitive attributes, providing insights for building fairer, GDPR and HIPAA-compliant ML models.

SKILLS

Programming Languages: Python, Swift, C++, Java, JavaScript, TypeScript, HTML, CSS, SQL

Tools and technologies: SwiftUI, Xcode, Unity Engine, Keras, MatPlotLib, Ski-kit Learn, Pandas, TensorFlow, MATLAB, Kafka, Apache Spark, PostgreSQL, NoSQL, Git, Jenkins, Docker, Hadoop, Distributed systems, REST APIs

PUBLICATIONS

WebDirect: A Web-based Project Repository | *IEEE Xplore Publication* *May 2023*

- The study presents a web-based project repository system using ReactJS, MongoDB, and cosine similarity for recommendations, addressing issues like complex UI and weak search filters.

A Web-based Project Repository using ReactJS | *IEEE Xplore Publication* *October 2022*

- The paper tackles common issues with project portals, introducing enhancements like improved search filters, project recommendations, trending sections, and user collaboration features.
- It advances the study of project portals, providing insights into functionality and usability, with potential for future adaptation to new technologies and business needs.