HARSHIT TIMMANAGOUDAR

+1 (858) 405-1729 ⋄ htimmanagoudar@ucsd.edu ⋄ linkedin.com/in/harshit-timmanagoudar

EDUCATION

University of California, San Diego, Master's of Science in Computer Science.　　　Sep 2024 - Jun 2026
Coursework: Data Systems for ML, Distributed Systems, Networked Services, Large Model Reasoning.

PES University, Bachelor of Technology in Computer Science and Engineering.　　　Aug 2020 - May 2024
Coursework: Data Structures, Advanced Algorithms, Graph Theory, Image Processing, Database Management.

SKILLS

Languages	C/C++, Python, Golang, Java, JavaScript, Bash, SQL, Neo4j
Frameworks/Tools	PyTorch, TensorFlow, Kubernetes, Docker, NodeJS, MongoDB, PostgreSQL, Git, Redis, AWS.
Soft Skills	Conflict Resolution, Growth Mindset, Resilience, Empathy, Constructive Feedback, Teamwork

EXPERIENCE

Graduate ML Systems Researcher　　　May 2025 – Present
University of California, San Diego - STABLE Labs　　　*La Jolla, CA*

- Solved memory fragmentation in **LLM KV cache systems** by architecting **dynamic memory allocation**, achieving **300x faster** store operations and superior scaling with only **11.5x degradation vs 40x degradation** at 16,000 tokens against HuggingFace's cache.
- Built **concurrent multi-user request** handling for KV Cache with atomic memory allocation, distributed sharding, and write-through buffering, maintaining consistent **sub-70 microsecond** write latencies across sequence lengths and **25,000 tokens/second** throughput.

AI Research Intern　　　Jan 2024 - Jun 2024
Peptris Technologies　　　*Bangalore, India*

- Spearheaded development of **data augmentation pipeline** for **Graph Attention Network**-based molecular conformation models, generating **25,000+** additional training samples and improving model accuracy by **12%**.
- Architected and optimized **sampling framework** for a **Diffusion Model**-based protein-ligand pose generation, collaborating with **ML engineering team** to reduce the generation time from **2.3s to 1.7s (35% improvement)**.

Machine Learning Intern　　　Jun 2023 - Sept 2023
Peptris Technologies　　　*Bangalore, India*

- Engineered automated **Large Language Model** inference pipelines for **protein sequence generation**, enabling **high-throughput analysis** of **500,000+** compound pairs with **92.5% accuracy**.
- Developed **3D U-Net** model architecture for **binding site detection** in voxel-based analysis, collaborating with **research scientists** to identify **30+** novel molecular interactions and accelerate **drug discovery** timelines.

PROJECTS

DistRAG. Tackled database scalability and accessibility issues by developing a **multi-node PostgreSQL cluster** with **RAG-powered natural language querying**, automated scaling, and intelligent caching for high availability and user-friendly data access. (GitHub)

StreamGrid. Addressed video distribution and fault tolerance challenges by developing a dynamic platform with **MPEG-DASH** streaming, gRPC cluster control, **consistent hashing** for content placement, and a **RAFT metadata layer** for resilient operations on AWS. (GitHub)

Git-LeetCode. Streamlined coding practice workflows by developing an automated **GitHub integration tool for LeetCode** that tracks submissions, synchronizes solutions, and provides **LLM-powered optimization insights**. (GitHub).

PUBLICATIONS

[1] **Enhancement of Malware Detection Systems using Mal-cGAN**.　　　Feb 2024
ISI, Springer Nature doi: 10.1007/978-981-97-2147-4_24