

PRINCE SAVSAVIYA

+1-951-377-3380 | princesavsaviya2023.learning@gmail.com | Riverside, CA | linkedin | github

Professional Summary

M.S. Computer Science candidate (UC Riverside, March 2026) with expertise in **ML systems, High-Performance Computing (HPC), and Distributed Systems**. Experience ranging from **optimizing GPU-accelerated physical simulations** to building scalable platforms like RetrieveX (+25% precision) and predictive Kubernetes autoscalers (-18% over-provisioning). Skilled in Python, C++, CUDA, and Cloud Infrastructure. Seeking SWE/ML role to design large-scale, high-efficiency systems.

Education

University of California, Riverside M.S. Computer Science, GPA 3.56/4.0	Riverside, CA	March 2026 (expected)
Adani Institute of Infrastructure Engineering, GTU B.E. Information & Communication Technology, GPA 3.6/4.0	Ahmedabad, IN	May 2024

Core Technical Skills

Languages & Query:	Python, C++, Java, C, SQL (PostgreSQL/MySQL), Bash, R
Systems & Cloud:	Docker, Kubernetes, AWS (EC2, S3, EKS), Linux, REST APIs, MPI
Data & Search:	Elasticsearch, Lucene, FAISS
AI, ML & Data Science:	Deep Learning (PyTorch, TensorFlow), scikit-learn, BERT/NLP, Computer Vision, MLOps (Optuna, Vector Search)
Performance & HPC:	CUDA, PETSc, DolfinX/FEniCSx, Gmsh
DevOps & Quality:	Git/GitHub Actions (CI/CD), Unit Testing, Monitoring, Hyper-parameter Tuning

Experience

Graduate Researcher — University of California, Riverside	Riverside, CA	May 2025 – Present
<ul style="list-style-type: none">Built a modular ALE-ready simulation framework integrating CFD and FEM solvers; validated against Turek-Hron benchmarks for high-fidelity FSI stability.Developed a parameterized Python+Gmsh benchmark generator with automated regression tests and logging, reducing setup time for new FSI/CFD cases by 80%.Optimized PETSc/CUDA kernels under GPU-aware MPI, reducing memory-transfer bottlenecks and improving multi-GPU throughput by 2×.Prototyping physics-informed ML predictors (e.g., Hamiltonian Neural Networks) to reduce fixed-point iterations in partitioned FSI coupling; designing feature sets and data pipelines to accelerate convergence.		
Research Intern — Adani University	Ahmedabad, IN	May 2023 – Dec 2023
<ul style="list-style-type: none">Built a physics-informed ML framework combining PDE ageing models with ensemble regressors (SVR, RF, GBR, DT, GPR), cutting MSE 98% and RMSE 86% vs. data-only baseline.Ingested 200,000+ battery-cycle records in Python; automated feature engineering and Optuna tuning, reducing experiment turnaround 30%.Publication: <i>Journal of Power Sources</i>, Vol. 627, 2025 (IF ≈ 7.9) — DOI 10.1016/j.jpowsour.2024.235771.		

Selected Projects

RetrieveX — Hybrid Search Engine	Django · Lucene · BERT · FAISS	Jan 2025 – Apr 2025
<ul style="list-style-type: none">Built medical information retrieval system (150,000 docs) merging BM25 keyword search with BERT+FAISS vectors, improving top-3 precision 25% and reducing latency 40%.Deployed as asynchronous REST API with Django; served P_{95} queries in <200 ms.		
Predictive Kubernetes Autoscaler	LSTM · Kubernetes · Helm · GitHub Actions	Jan 2025 – Apr 2025
<ul style="list-style-type: none">Developed an LSTM-driven autoscaler forecasting CPU/memory load to adjust replicas dynamically, reducing over-provisioning 18%.Deployed to AWS EKS with Helm and CI/CD pipelines; sustained 99.9% SLA under 4× load spikes.		
Risk-Analytics Microservice	FastAPI · Pandas · PostgreSQL · Docker · AWS	Sep 2024 – Jan 2025
<ul style="list-style-type: none">Designed real-time analytics service ingesting multi-asset data (equities, bonds, FX) with vectorized pipelines; delivered metrics in <150 ms.Automated deployment (CI/CD) with GitHub Actions → AWS ECR → EC2; 99.9% uptime with Prometheus–Grafana monitoring.		