# Varun Phanindra Shrivathsa

312-478-2342 | vphan@uic.edu | Chicago, IL | linkedin.com/in/varun-p-shrivathsa
www.varun-p.com | github.com/varunpshrivathsa | medium.com/@varunpshrivathsa

## Summary

UIC MSCS Graduate with 2 years of combined internship and research experience in machine learning and software development. Skilled in Agentic AI, Deep Learning with PyTorch/TensorFlow and scalable distributed architectures. Experienced in building LLM applications and deploying microservices on AWS using Docker and Kubernetes.

## Skills

**Programming Languages:** Python, C/C++, SQL, JavaScript
**ML & DL:** PyTorch, TensorFlow, Keras, SparkML, Scikit-learn, XGBoost, Hugging Face, MLFlow, OpenCV
**Generative AI:** LLMs (GPT, LLaMA, Mistral), LangChain, LoRA/QLoRA
**Frameworks & Systems:** FastAPI, Ray Tune, Flask, React, Databricks, gRPC, Kafka, Airflow, Spark, Redis, Linux
**Data Engineering & Databases:** PostgreSQL, MongoDB, DynamoDB, Pinecone, FAISS
**MLOps & Deployment:** MLflow, Docker, Kubernetes, AWS, GCP, Terraform, GitHub Actions
**Monitoring & Testing:** Prometheus, Grafana, PyTest
**GPU Acceleration:** CUDA, TensorRT, cuDNN

## Experience

**Machine Learning Engineer Intern** — Jan 2025 – Present
*G19 Studio* — *Chicago, USA (Remote)*

- Worked with G19 Studio through UIC's Advanced NLP course and continued as an extended internship.
- Developed 'TwinVerse' a human digital twin platform for stress mitigation using real-time wearable sensor data.
- Built TCN and PPO-RL agents for physiological signal forecasting achieving 25% faster simulation convergence.

**Software Developer - ML Intern** — Jan 2024 – Apr 2024
*Mekhalyn* — *Bangalore, India*

- Built a recruiter analytics platform using FastAPI, React and PostgreSQL, orchestrated via Kubernetes.
- Developed a LoRA-tuned LLM pipeline with FAISS-based semantic retrieval to summarize 10k+ resumes.
- Optimized API throughput and inference latency, achieving 28% lower compute cost under production workloads.

**Research Intern** — Jun 2023 – Jun 2024
*Indian Institute of Science (IISc)* — *Bangalore, India*

- Led the development of GIS image processing for environmental impact analysis of Bangalore STRR project.
- Built a multi-spectral CNN achieving 92% segmentation accuracy for land-cover classification.
- Awarded KSCST Research Funding to extend geospatial validation and automated GIS pipeline optimization.

## Projects

**GenCost: Adaptive Multi-Agent LLM Cost Optimization Platform**

- Designed a multi-agent system to route prompts across LLMs based on real-time cost, latency and quality metrics.
- Integrated contextual bandit and PPO to achieve 40% cost reduction while maintaining response quality.
- Deployed FastAPI backend with PostgreSQL, Redis on AWS ECS (Fargate) with CI/CD achieving 99.9% uptime.

**VideoTune: Multimodal Video Recommendation System**

- Developed a cross-modal transformer using CLIP, Wav2Vec2 and BERT embeddings for engagement prediction.
- Optimized multi-task learning via dynamic loss weighting and Pareto frontier tuning, boosting accuracy by 22%.
- Containerized and deployed real-time stack on AWS ECS with S3, Faiss indexing, and Prometheus monitoring.

**DistFlow: Distributed Task Scheduling and Auto-Scaling Engine**

- Built task scheduler with DAG execution engine, achieving 10K+ tasks/day and 45% faster workflow completion
- Implemented fault-tolerant execution with exponential retries and circuit breakers for 99.95% reliability.
- Deployed on AWS EKS using Redis, and PostgreSQL with Grafana dashboards for real-time insights.

## Education

**University of Illinois Chicago (UIC)** — Chicago, USA
*Master of Science in Computer Science* — *Present – Apr 2026*

- Thesis: Robotic Planning, Localization & Exploration (Advisor: Prof. Wenhao Luo).

**Dayananda Sagar University** — Bangalore, India
*Bachelor of Technology in Computer Science and Engineering* — *Jun 2020 – Jun 2024*

## Publications

"Environmental Impact Analysis using Satellite Image Processing: Bangalore STRR." *2024 IEEE - Published*
"Deepfake Detection using LSTM and XResNet." *IJRASET-Published*