

HARDIK SRIVASTAVA

+1 (206) 773-9806 | hardiksrwork@gmail.com | [Portfolio](#) | [LinkedIn](#) | [GitHub](#) | [ResearchGate](#) | [Google Scholar](#) | [LeetCode](#)

EDUCATION

University of Washington

Master of Science in Data Science (GPA - 3.9/4)

Seattle, WA

Sept 2025 - Mar 2027

SRM Institute of Science and Technology

Bachelor of Technology in Computer Science with specialization in Big Data Analytics (GPA - 3.82/4)

Chennai, TN

May 2019 - Jun 2023

WORK EXPERIENCE

Graduate Research Assistant (P.I. - Prof. Zaid Harchaoui)

Paul G. Allen School of Computer Science, University of Washington

Sept 2025 - Present

Seattle, WA

- Architected an **authorship obfuscation** system by fine-tuning 16+ modular LoRA adapters (Qwen2.5-32B); built **metric-driven agents** to guide controllable style transfer via learned style embeddings and constrained decoding, achieving **40%** privacy improvement.
- Developed a **gradient-free few-shot optimizer** using in-context learning to dynamically weight 16-adapter ensembles, reducing author-specific configuration time from days to minutes. Outperformed grid-search baselines by **23%** in privacy-utility Pareto efficiency.
- Productionized the **agentic workflow** on **HPC clusters** utilizing Mistral-24B for robust evaluation and policy refinement. Engineered multi-objective feedback loops to enable **real-time policy iteration**, maintaining **88%** semantic similarity and **5%** perplexity drift.

Applied Scientist 2

JPMorganChase & Co.

Jun 2023 - Aug 2025

Hyderabad, TS

- Developed a custom **cost-sensitive Focal Loss objective** for XGBoost & CatBoost to penalize false negatives; retrained production models to secure a **6%** recall boost and **\$220M+** in annual savings while utilizing TreeSHAP to maintain strict feature interpretability.
- Productionized semi-supervised Graph Neural Networks across 60M+ transaction nodes, introducing **topology-aware edge sampling** and dynamic **neighborhood aggregation** to scale **fraud-detection** to 2M+ transactions/day with a **9.2%** precision boost.
- Architected a GPU-served **contrastive multimodal Transformer** (ViT + BERT-style decoder) for OCR-based check fraud detection, serving **500K+** requests/day. Improved fraud precision by **19%** while reducing latency via batched mixed-precision inference.
- Built a **Learning-to-Rank** information retrieval system over **3B+** KYC records using Pairwise Ranking loss to flag fraud entities; Outperformed the existing TF-IDF baselines, cutting **62% false positives** and saving **700+** analyst review hours daily in production.
- Engineered highly-scalable ML platforms and **distributed ETL pipelines** in AWS Glue for **terabyte-scale** financial data. Developed automated feature selection workflows that boosted downstream fraud capture by **14%**, reducing manual feature engineering overhead.

Research Intern

McGill University - *Mitacs Scholar*

Jun 2022 - Sept 2022

Montreal, QC

- Engineered a dense retrieval system for predictive **vocabulary generation** by training DeBERTa to encode image captions and vocab into a shared latent space; implemented **FAISS** for efficient k-nearest neighbor retrieval, outperforming the baseline by **6.9 MAP**.

Research Intern

Samsung Research

Jan 2022 - Jun 2022

Bangalore, KN

- Optimized production-scale **NLU intent routing** for Samsung Bixby via neural graph matching, reducing false positives by **12%**; Deployed INT8 inference via **Quantization-Aware** training for word disambiguation models, surpassing FP32 baselines by **2.9% F1**.

SELECTED PUBLICATIONS

- [Srivastava, H.](#) **CM-BERT: Multi-Modal Sentiment Analysis Using Text & Audio for Customer Support Centers** (2023)
- [Srivastava, H.](#) **StyleLM: Neural Text Style Transfer with Custom Language Styles for Personalized Communication Systems** (2023)
- [Srivastava, H.](#) Using NLP Techniques for Enhancing Augmentative and Alternative Communication Applications (2021)

PROJECTS

Agentic Ad Intelligence Engine [\[code\]](#) Architected a **multi-agent orchestration** pipeline for X (Twitter) using **Grok 4.1** to generate 100+ dynamic user personas and personalized ad copy, achieving **sub-5-second** end-to-end latency. Engineered a dynamic routing mechanism that ranked and delivered ads based on live engagement feedback, which increased click-through rates by **16%** in pilot campaigns

KeyCognition: Behavioral AI Platform [\[code\]](#) Processed **136M+** keystroke events in Databricks to predict a user's cognitive load by analyzing micro-temporal typing anomalies, such as backspaces and key flight times. Trained a hybrid **behavioral classification** model fusing LSTMs with K-Means clustering to group and extract these behavioral patterns into latent embeddings.

SKILLS

CS/Tooling	C++, Python, Java, SQL, PyTorch, TensorFlow, Sklearn, Spark, CUDA, LangGraph, HuggingFace
AI/Research	Machine Learning, Natural Language Processing, Generative AI, Large Language Models, Ranking, Agentic Workflows Time Series Forecasting
Infra/MLOps	AWS, Kubernetes, LangChain, MLFlow, RAG, Distributed Systems, Feature Stores, Vector Search