

# HARDIK SRIVASTAVA

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

## EDUCATION

<b>University of Washington</b> <i>Master of Science in Data Science</i>	Seattle, WA Sept 2025 - Mar 2027
<b>SRM Institute of Science and Technology</b> <i>B.Tech in Computer Science Engineering with specialization in Big Data Analytics (GPA - 3.82/4)</i>	Chennai, TN May 2019 - Jun 2023

## WORK EXPERIENCE

<b>Graduate Research Assistant (P.I. - Prof. Zaid Harchaoui)</b> Paul G. Allen School of Computer Science, University of Washington	Sept 2025 - Present <i>Seattle, WA</i>
--	---

- Engineered and fine-tuned **modular LoRA adapters** for **controllable rewriting**, then built a **metric-driven agentic system** to plan and execute authorial style remixing via learned style representations and attribution classifiers under privacy-fidelity constraints.
- Formalized **16+ authorial style axes** as interpretable control variables with **closed-loop logic** for dynamic perturbation scaling. Improved anonymity by 40% on AuthorMix (30K texts/14 authors) while preserving 88% semantic similarity and 5% perplexity drift.
- Optimized agent policies via **multi-objective feedback** and conducted empirical analysis across multilingual datasets. Benchmarked the system against larger LLM baselines, achieving **privacy gains** at **3x lower inference cost** and **34% fewer generated tokens**.

<b>Applied Scientist 2</b> JPMorgan Chase & Co.	Jun 2023 - Aug 2025 <i>Hyderabad, TS</i>
--	---

- Led **fraud detection** initiatives saving over **\$220M annually** by developing high-recall classifiers using **XGBoost** and **CatBoost**; leveraged **multi-objective optimization** to preserve feature interpretability for business partners.
- Deployed Semi-Supervised Graph Neural Networks on 60M+ transaction nodes; introduced **topology-aware edge sampling** and **dynamic neighborhood aggregation**, boosting fraud-detection **precision** by **9.2% over 2M transactions/day**.
- Designed and productionized a CLIP-style **Contrastive Pre-Trained Transformer** (ViT encoder & BERT-style decoder) for **Document OCR** to detect Check fraud, processing **500K+** checks daily with a **19% boost** in text matching accuracy.
- Built **Learning-to-Rank** entity matching system for KYC using Pairwise Ranking loss to alert fraud entities; Outperformed the TF-IDF baseline on **3B+ records**, cutting **62% false positives** and saving **700+** manual review hours daily.
- Developed data pipelines for high-volume transactions in PySpark. Performed **optimal feature selection** using **Decision Trees** that boosted **Fraud Capture rate** by **14%** and enhanced predictive accuracy for risk scoring models.

<b>Research Intern</b> McGill University - <i>Mitacs Scholar</i>	Jun 2022 - Sept 2022 <i>Montreal, QC</i>
---	---

- Developed **RoBERTa-based sentence embeddings** on AAC image captions to capture word-context. Achieved **Krippendorff's** alpha reliability score of **0.82** for context-aware vocab retrieval; deployed the model in **ClickAAC** mobile application.

<b>Machine Learning Engineer - Intern</b> Samsung Research	Jan 2022 - Jun 2022 <i>Bangalore, KN</i>
---	---

- Optimized **word-context relationships** for Samsung Bixby search using bipartite graphs, reducing **false-positive** preds by **12%**. Matched **FP32** perf for INT8 inference using quantization-aware training for word-disambiguation task with a **2.9% F1** gain.

## PROJECTS

**Agentic Ad Intelligence Engine** [\[code\]](#) Built a hyperpersonalized ad intelligence system that constructs **dynamic user personas** from behavioral signals and generates **context-aware ad variants** conditioned on audience attributes; implemented **feedback-driven ranking** and **attention-aware decision logic** using ensemble engagement predictions for **smart ad delivery** enabling scalable personalization.

## PUBLICATIONS

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

## ACCOMPLISHMENTS

- Certifications** — AWS Certified Machine Learning Specialist, AWS Certified Cloud Practitioner
- Hackathons** — xAI Hackathon'25 (**Bronze**), UWxDatabricks Hack'25 (**Bronze**), JPMC AWS DeepRacer'23 (**Gold**)

## SKILLS

<b>AI/Research</b>	Machine Learning, NLP, Sequence Modeling, Statistical Modeling, Probabilistic Reasoning, Graph-ML, Transformers
<b>CS/Tooling</b>	Semi-Supervised Learning, Ranking, Agentic-AI, LLM Inferencing, Time Series Forecasting, Data Analysis
<b>Infra/MLOps</b>	C++, Python, SQL, PyTorch, TensorFlow, Caffe, Sklearn, Spark, Kafka, W&B, HuggingFace, Optuna, Git AWS, Docker, Kubernetes, Flask, MLFlow, NVIDIA Triton, Feature Stores, Data Version Control, Jenkins