9179401695 **Atul Dhingra** dhingra.atul92@gmail.com
Milpitas, CA linkedin.com/in/dhingratul/

Summary

Machine Learning leader with 9+ years of experience (3+ years in leadership) across FinTech, Autonomous Checkout, Autonomous Vehicles, and Digital Health. Proven track record of designing and scaling applied ML systems using LLMs, Computer Vision, Reinforcement Learning, Deep Learning, and Edge AI. Skilled at translating research into production across cloud, edge, and on-prem compute to deliver measurable business impact. Hold multiple patents and publications in Biometrics and Computer Vision, with work featured in peer-reviewed conferences and journals. Key contributions include:

- Leading PayPal's Agentic AI initiative, leveraging LLMs and RL with verifiable rewards to build multi-turn agents capable of long-horizon decision-making.
- Scaled Standard AI's autonomous checkout from 2 → 40+ stores, deploying CV models across in-store clusters, cloud-native, and edge, saving $4M+ annually.
- Designed the early Edge compute framework at NIO, which became the foundation for the Adam supercomputer architecture later deployed in 100K+ L4 vehicles.

Experience

Staff Machine Learning Engineer, PayPal, San Jose, CA **Jul '23 - Present**

- Leading PayPal's Agentic AI initiative, leveraging LLMs and Reinforcement Learning with verifiable rewards to build multi-turn conversational agents for long-horizon decision-making.
- Trained multi-billion parameter LLMs with PPO-style objectives and KL regularization, scaling rollouts across 100+ environments; leveraged distributed training (FSDP on H100/B200 clusters) to enable applied Agentic AI models for production readiness.
- Driving the transition from R&D to production by ensuring agentic models meet inference latency, quantization, and SLA requirements, with robust API/tool integration.
- Improved ML developer productivity by 30% by deploying LLM-based coding and workflow assistants, informing build-vs-buy decisions and shaping PayPal's internal ML strategy.
- Delivered end-to-end applied coding LLMs including dataset curation, finetuning, RAG integration, and evaluation, that powered developer-facing ML tools and accelerated prototyping.

Engineering Manager, Machine Learning, Standard AI, San Francisco, CA **Nov '19 - May '23**

- Scaled autonomous checkout from 2 → 40+ stores across 3 hardware platforms, handling 5× greater environment complexity and foot traffic, saving $4M+ annually.
- Designed next-gen edge-based perception hardware, delivered in 6 months, improving store-level margins by $1M annually.
- Launched new product lines in shopper trajectory analytics and visual re-identification, boosting profit margins and user retention by 15%.
- Built an automatic labeling system for long tail data distribution detection in perception models, enabling 10× faster software deployment and reducing dev costs by 85% ($1M YoY savings).
- Automated training and deployment workflows for CV models, improving accuracy by 25% while cutting operational costs by 90% ($1M YoY savings).
- Streamlined CV/DNN deployment by building reliable inference workflows on Vertex AI, accelerating production adoption for 5 internal teams.
- Built and scaled a high-performing global ML team of 7 engineers and researchers—leading hiring, onboarding, and mentorship while delivering cross-functional, high-impact ML projects under tight deadlines
- Defined team roadmap and success metrics that aligned perception and analytics R&D with product goals, accelerating time-to-deployment for autonomous checkout.

Perception Engineer, NIO, San Jose, CA **March '18 - Nov '19**

- Designed the early edge compute framework that became the foundation for NIO's Adam supercomputer architecture, later deployed in 100K+ L4 autonomous vehicles.
- Optimized lidar-based scene understanding models (two-stage detection with BEV + 3D voxels), achieving compressed and accelerated performance for production.
- Improved edge inference efficiency by 5× with ¡1% accuracy trade-off, enabling real-time AV perception on embedded hardware
- Developed a low-level inference libraries with custom ops for hardware accelerators, supporting concurrent deployment of multiple DNN models.

- Mentored senior engineers and interns in deploying DNN models on embedded platforms.

Machine Learning Engineer, Otsuka Digital Health, Princeton, NJ **July '17 - March '18**

- Boosted medical claims cost prediction accuracy by 90% with LSTM-based models, enabling better patient prioritization and care
- Designed and implemented advanced deep learning and ML algorithms for predictive modeling, along with novel methods to describe and visualize medical claims data.

Graduate Research Assistant, Rutgers University, New Brunswick, NJ **July '16- June '17**

- Developed a novel algorithm for face clustering based on multiple facial attributes
- Developed and published a novel face recognition algorithm for aggregating visual features based on clustering in a multi-shot video-to-gallery template retrieval problem in an unconstrained environment
- Investigated the role of face data and attribute bias in automated photo-sketch generation

Visiting Researcher, Indian Institute of Technology, Delhi, India **Dec '11 - July '15**

- Developed and published a robust speaker verification algorithm invariant to noise and multi-channel input using GFCC, MFCC and i-vectors.
- Developed and published novel decision tree based method for error analysis for eye movement tracking for biometrics.
- Independently led data collection, curation, labeling and management of human irises datasets from 50+ users in a period of one month.

Publications

Atul Dhingra, G Sood, " Scaling ML Products At Startups: A Practitioner's Guide", arXiv:2304.10660

Atul Dhingra, G Sood, "Indian Electoral Roll Corpus", Harvard Dataverse, doi.org/10.7910/DVN/OG47IV

Atul Dhingra, M. Jeevan, M. Hanmandlu, B.K Panigrahi , "Robust Speaker Verification using GFCC based i-vectors", in Proceedings of the IEEE International Conference on Signal, Networks, Computing, and Systems 2016 (Springer)

Atul Dhingra, K Vishal, "Wielding Audio-Books for Visually Impaired Using Gesture Recognition", International Journal Of Advanced Research Trends In Engineering And Technology; 2(5), pp. 64-68, 2015

Atul Dhingra, A Kumar, M. Hanmandlu, B.K Panigrahi , "Biometric Based Personal Authentication Using Eye Movement Tracking", SEMCCO 2013, Part II, LNCS(Springer) 8298, pp. 248-256, 2013

Patents

Atul Dhingra, et al, "One or more cameras for use in an autonomous checkout in a cashier-less shopping store and otherwise"", US Application No. 18/539,228, December 13, 2023

Atul Dhingra, et al, "Subject-tracking in a cashier-less shopping store for autonomous checkout for improving item and shelf placement and for performing spatial analytics using spatial data and the subject-tracking", U.S. Application No. 18/522,104, November 28, 2023

Atul Dhingra, et al, "Systems and methods for performing spatial analytics using spatial data related to a cashier-less shopping store for autonomous checkout", U.S. Application No.: 63/428,373, November 28, 2022

Atul Dhingra, et al, "Machine learning-based re-identification of shoppers in a cashier-less store for autonomous checkout ",U.S. Application No. 17/988,650, November 16, 2022

Skills

Programming Languages: Python

Tools/ APIs: Torchtune, TorchRL, Langchain, Hugging Face, Milvus, Pytorch, Tensorflow, TensorRT, Vertex AI, MLOps, NLP, Edge Computing, Numpy, Git, Keras, OpenCV, Dlib, Pandas, Selenium, VLFeat, Unity3D, Data Analysis, Linux

Edge Hardware: Nvidia Drive AGX, Nvidia Jetson, Intel Accelerators, Others (Under NDA)

Softwares: Scrum, Kanban, Jira, Smartsheets, github, dvc, GCP, databricks, looker, Pager Duty

Education

M.S in Computer Science, Rutgers University, NJ, USA

Visiting Researcher at Biometrics Research Lab, Indian Institute of Technology, Delhi, India

B.E in Instrumentation & Control Engineering, University of Delhi, India