

ABHIRAM SANJAY DHARME

abhiram.cse.iitd@gmail.com • +91 70585 00898 • abhiramdhar.me • github.com/abhiramDhar.me • linkedin.com/in/abhiram-dhar.me

EDUCATION

Indian Institute of Technology Delhi

Dual Degree (B.Tech + M.Tech), Computer Science & Engineering

Expected: May 2027

CGPA: 8.03/10.00

EXPERIENCE

Google

May 2026 – Jul 2026

Software Engineering Intern • Gemini Cloud Assist / Google Cloud AI Diagnostics and Monitoring • Bangalore, India

- Designed and implemented an **MCP endpoint** enabling **Gemini Cloud Assist** to auto-diagnose and resubmit failed enterprise batch jobs (e.g., OOM failures), replacing manual log-debugging with **one-click resolution** — a step toward **agentic, self-healing cloud operations**.
- Migrated a monitoring library from **Cloud Monarch** to Google's company-wide **GMon** framework, reducing internal tech debt and improving production-metric observability across legacy systems.

Coinbase

May 2025 – Jul 2025

Machine Learning Engineering Intern • CBGPT Team • Bangalore, India

- Built a parameterizable, production-ready knowledge-gap classification pipeline to analyze chatbot failure states, processing **36,000+ real user queries** to isolate an **18% support documentation gap**.
- Designed a highly **concurrent FAISS-based retrieval system** over large-scale article knowledge bases; implemented **multi-threaded topic assignment** and **asynchronous LLM batching** to drastically reduce pipeline execution time.

Torch Investments

Jan 2024 – Mar 2024

Machine Learning Engineering Intern • New Delhi, India

- Developed a **Python modeling and backtesting pipeline** using **LightGBM/CatBoost** for Fortune-500 equity-return prediction; a 2014–2024 backtest yielded **21% CAGR**, backing a live **\$3M+ strategy**.

SELECTED ML / SYSTEMS PROJECTS

CycloFormer (Master's Thesis)

Aug 2025 – Present

PyTorch, Transformers, Wearable Biosignals, Scaling Laws; w/ Google DeepMind

- Designed **CycloFormer**, a $\mathbb{Z}_{16}$ -invariant **Transformer** for wrist-sEMG hand-pose estimation, composing channel-shared TDS-CNNs, circular RoPE, and permutation-invariant attention pooling to bake in exact wristband-rotation invariance by construction; manuscript submitted to **NeurIPS 2026**.
- Achieved **state of the art** on emg2pose: a 4M-parameter model beat the strongest 6M baseline on every generalization split with **33% fewer parameters**; scaling to 48M widened the Stage-split margin by **2.5° / 3.3 mm PA-MPJPE**.
- Ran the **first controlled sEMG-vs-egocentric-vision comparison** under fingertip self-occlusion; identified a consistent crossover where sEMG becomes the more reliable modality beyond **two occluded fingertips**, positioning biosignals as complementary to vision for dexterous manipulation.
- Derived a **data-scarcity-aware scaling law** $L(N, D) = \epsilon_{\infty} + aN^{-\alpha} + bD^{-\beta} + c(N/D)^{\gamma}$ that recovers Chinchilla in the data-rich limit and captures data-limited overfitting; fit over 8 model sizes and 5 data fractions, explaining **98.8%** of PA-MPJPE variance and projecting **~34× more session-hours** to close the gap to vision at compute-optimal scale.

Representation Learning, VLMs, and Diffusion • PyTorch, ViT, Transformers, Diffusion

- Implemented **CLIP** from scratch: ViT-S/16 image encoder, causal-masked text Transformer, dual-projection contrastive head (CLS + GAP with auxiliary loss), DropPath regularization, and bicubic positional-embedding interpolation for higher-resolution probing.
- Trained **DINO** self-supervised distillation with multi-crop augmentation and centring (no text supervision); built a two-stage vision-language model with image-text alignment followed by chain-of-thought QA with explicit numerical-stability treatment of the CoT logits.
- Trained a **VAE + Latent Diffusion Model** end-to-end for generative modeling; benchmarked representation quality across CLIP, DINO, and the VLM via linear probes, t-SNE clustering, and cross-modal retrieval.

RRM Plus: Client-Aware WiFi Resource Management • GNNs, Safe RL, Network Simulation

- Generated a **1,000-scenario / 4,000-AP / 24,000-client** WiFi training corpus and modeled dense wireless networks as heterogeneous graphs over APs, clients, and interference edges.
- Trained a safety-constrained GNN/RL controller with trust-region checks, QoE safety critic, per-AP change budgets, and rollback guards for edge-throughput regressions.

Visual Question Answering • PyTorch, ResNet-101, Transformers, Cross-Attention

- Built a multimodal VQA model using a frozen ResNet-101 image encoder, Transformer text encoder, and multi-head cross-attention from the question token to visual features.
- Used progressive training from frozen backbone to fine-tuned image encoder, with held-out question-type evaluation to probe compositional generalization.

Neural Machine Translation with Cross-Lingual Transfer • *BERT, BiLSTM, Attention*

- Implemented four English-to-Hindi NMT variants with GloVe/BiLSTM, Bahdanau attention, frozen-BERT encoders, and fine-tuned-BERT encoders; used label smoothing, beam search, and teacher-forcing schedules.
- Adapted the Hindi model to Hindi-to-Marathi transfer, exploiting shared Devanagari script and morphology for low-resource cross-lingual generalization.

ResNet-18 from Scratch with Seven Normalization Schemes • *PyTorch, Computer Vision*

- Implemented **ResNet-18** and every normalization variant from scratch as `nn.Module` subclasses (Batch, Instance, Batch-Instance, Layer, Group, custom BN, no-norm baseline); trained on a 100-class ImageNet subset with SGD + cosine annealing, MixUp ($\alpha=0.4$), RandAugment, label smoothing, and mixed precision for 100 epochs across all variants.
- Sanity-checked the custom BN against PyTorch's built-in (**84.2% vs 84.5%**; gap traced to Bessel-corrected variance); ran **Grad-CAM** analysis across visually similar classes (snakes, reptiles) to interrogate what each normalization scheme actually learned to attend to.

Parallel Matrix Modification on CUDA • *CUDA, C++, Parallel Algorithms*

- Implemented a GPU pipeline using counting sort, Blelloch prefix scan, and binary search; used pinned host memory and CUDA streams to overlap transfers with kernel execution.

Kernel Extensions in xv6 • *C, Operating Systems, Systems Programming*

- Added five kernel subsystems to xv6: login authentication, syscall access control, custom interrupt modes, priority-boosted scheduling, and disk-backed page swapping.

Reliable Transport over UDP • *C, TCP Reno / CUBIC, Mininet, Ryu*

- Built a reliable file-transfer protocol on UDP with application-layer ACKs, retransmission, and **TCP Reno + CUBIC** congestion control; benchmarked on Mininet across packet-loss and RTT sweeps.

Cache Hierarchy Simulator • *C++, OOP, Computer Architecture*

- Cycle-accurate simulator supporting direct-mapped, set-associative, and fully-associative caches with LRU/FIFO and Write-Through/Back $\times$ Allocate/No-Allocate policies; miss-penalty modelling.

SKILLS

Languages: Python, C++, C, Java, JavaScript, OCaml, LaTeX

ML/Data: PyTorch, TensorFlow, NumPy, Pandas, scikit-learn, FAISS, LightGBM, CatBoost

Methods: Transformers, LLM retrieval/ranking, contrastive learning, diffusion models, GNNs, reinforcement learning

Systems: Linux, Git, CUDA, xv6, sockets, Mininet, Ryu/OpenFlow, computer architecture

ACHIEVEMENTS

- **JEE Advanced 2022:** All India Rank 1055 among $\sim 250,000$ candidates; JEE Main 2022: AIR 1342 among $\sim 900,000$ candidates.
- **Codeforces Expert** (peak rating 1672); Tower Research Capital Data Science Challenge top-50 institute-wide team, 2024 and 2025.
- **Founder, ARIES** – Artificial Intelligence Society, IIT Delhi; organized ML and generative-AI workshops, panels, and hackathons.