YASH BHARDWAJ

yash.bhardwaj@polytechnique.edu | +33 745764696 | LinkedIn | GitHub | Homepage

SUMMARY

- MSc AI @ École Polytechnique; published at KDD'25 (Oral) and ICCV'25 on multimodal video understanding.
- Specialize in end-to-end multimodal research: dataset/annotation design → evaluation harnesses → VLM/MLLM systems (fusion, structured extraction, grounded QA/RAG).
- Seeking Research Intern (Apr–Aug 2026) in multimodal learning, diffusion, and LLM/VLM pre- or post-training.

EDUCATION

École Polytechnique   *Sept 2025 – Present*
Masters of Science and Technology in Trustworthy and Responsible AI   *Paris, France*
Courses(MS): Machine Learning, Deep Learning, Computer Vision, Collaborative Learning, Reinforcement Learning

Birla Institute of Technology And Science (BITS), Pilani   *Aug 2017 – Aug 2021*
Bachelors of Engineering in Computer Science   *Pilani, India*

SELECTED PUBLICATIONS

- VideoConviction: A Multimodal Benchmark for Human Conviction and Stock Market Recommendations – **KDD'25 Oral**
- FinCap: Topic-Aligned Captions for Short-Form Financial YouTube Videos – **ICCV'25 Poster**

WORK EXPERIENCE

Georgia Institute of Technology | Research Intern   Atlanta, Georgia | May 2024 – Sep 2025

- **VideoConviction:** Built a multimodal benchmark (288 videos, 6K+ annotations, 457 hours) to study conviction cues in financial advice, benchmarking transformer-based MLLMs (Gemini 2.5, GPT-4o, Qwen 2.5).
- **FinCap**: Developed a short-form financial video captioning benchmark (624 clips, 7 modality setups, 5 tasks), benchmarking modality fusion strategies for recommendation, visual features, sentiment, and entity extraction.

Urban Company | Software Developer 2 (Promoted in 2023)   Bangalore, India | Aug 2021 – Mar 2025

- **Product Catalog Cache:** Designed a distributed cache datastore for 27K products across 5 countries using Kafka, Redis, and MongoDB; reduced service latency by 40% and supported recommendation pipelines.
- **Dynamic Pricing Framework:** Deployed SARIMA + XGBoost–based time-series models for adaptive pricing by demand and region, improving revenue by 4% on $80M+ transactions.

IIIT Delhi | Research Intern   New Delhi, India | Jan 2021 – June 2021

- **Author Profiling:** Modeled authorship patterns in scientific publications using GNNs (GNN-XML) and AnnexML (KNN) on the S2ORC corpus (600K papers, 160K authors), achieving P@1: 23%, P@3: 15%.

Bridge Your Network | Software Engineering Intern   Dublin, Ireland | Jan 2021 – June 2021

- **App Development:** Built React interfaces and Ruby on Rails APIs for a networking platform (40M+ users), enhancing profile features and backend reliability.

SELECTED PROJECTS

- **Fast Diffusion via LoRA Distillation:** Distilled SD1.5 into a LoRA student for 2–8 step text-to-image generation (up to 18× faster); built a CUDA-timed benchmark with CLIP alignment + qualitative grids.
- **Emotional Scene Generation:** Fine-tuned Stable Diffusion v1.5 for 8 emotion classes on 100K images using LoRA (0.4% trainable params); achieved 15× faster training and 80% memory reduction vs full fine-tuning.
- **MedCLIP Mini:** Compact CLIP-style dual-encoder (ResNet-18 + DistilBERT) trained with contrastive (InfoNCE) loss on ROCO dataset achieving R@1 28% / R@10 55% with FAISS image–text retrieval.
- **Multimodal RAG:** Building a vision-based RAG system for scientific papers using PaliGemma2-3B with equation-aware retrieval to improve accuracy on math-heavy queries over traditional OCR baselines. *(Present)*

KEY ACCOMPLISHMENTS AND EXPERIENCES

- **Charpak Masters Scholarship:** 56 selected/2500+ applicants – French Ministry of Europe & Foreign Affairs.
- **Competitive Programming:** World rank 4, 6 in codechef long challenge Sept'22, Oct'22 (10K+ participants).

SKILLS

Programming Languages	Python, C++, JavaScript, C
Frameworks/Libraries	PyTorch, TensorFlow, Hugging Face, scikit-learn, NumPy, Pandas
AI/ML Areas	Multimodal Learning, LLMs, VLMs, Transformers, Diffusion Models, NLP, RAG
Tools	Git, Docker, Snowflake, MySQL, Linux, Kafka, Redis, MongoDB, AWS