Arpon Kapuria

Email: arpkapuria@gmail.com **Mobile:** +880 1797 288 296
Socials: LinkedIn | GitHub | LeetCode **Website:** arpon-kapuria.github.io

Summary

I'm a Computer Science graduate, deeply passionate about research and building AI systems that improve our day-to-day life. Alongside my technical pursuits, I enjoy teaching as a way to give back and simplifying complex ideas for others. Outside work, I love to travel, write and explore new places.

Research Interests

Deep Learning, Representation Learning, LLM Reasoning, RAG

Education

National Institute of Technology Tiruchirappalli Tiruchirappalli, India
B.Tech in Computer Science & Engineering **December 2020 – June 2024**

Thesis: Thresholding Binarized Neural Networks to Improve Accuracy in LLM Training
Courseworks: Data Structures and Algorithms · Operating Systems · Database Management Systems · Computer Networks · Machine Learning · Artificial Intelligence · NLP · Image Processing · Augmented and Virtual Reality · Technical Writing

Work Experience

Advanced Machine Intelligence Research Lab Dhaka, Bangladesh
Research Intern | *Advisor: Prof. M. Firoz Mridha* **February 2025 – Present**

- Investigate LLM reasoning capabilities to improve performance on complex tasks, with a focus on reducing hallucinations and applications of Retrieval Augmented Generation in Medical AI.

Indian Institute of Technology Bombay Mumbai, India
Research Intern | MeDAL Lab | *Advisor: Prof. Amit Sethi* **May 2023 – July 2023**

Project 1: Enhancing Self Supervised Learning framework - BYOL

- Conducted an in-depth survey on self-supervised learning, focusing on contrastive learning frameworks and their core principles.
- Reproduced BYOL and simCLR in PyTorch to establish a baseline for improvements.
- Introduced changes to BYOL by incorporating a novel loss function, algorithmic modifications and architectural adjustments to improve representation learning.

Project 2: Radiology DICOM Image Anonymization

- Created a user-friendly Flask-based system for DICOM image processing and anonymization using Pydicom, ensuring HIPAA-compliant privacy across **5,000+ medical scans**.
- Implemented and evaluated the CRAFT model for text detection in X-ray images and extracting clinically relevant annotations.

National Institute of Technology Tiruchirappalli Tiruchirappalli, India
UG Research Assistant | Industrial Automation Lab **October 2022 – February 2023**

Project: Assam Smart Home Automation *Advisors: Prof. M. Brindha, Prof. G. S. Ilango*

- Developed a Flutter app to automate energy operations and digital billing for a solar-powered microgrid in Assam, India, serving **200+ households**.
- Integrated Flutter frontend with Django-based REST API and deployed an MQTT broker enabling real-time data exchange with **50+ IoT sensors** for energy consumption and control.
- Project supported by the Ministry of Science & Technology, Government of India, under the initiative **SUSTENANCE** to achieve Novel Carbon Neutral Energy Communities.

Skills

- **Programming Languages:** C, C++, Python, Dart, JavaScript, SQL
- **Frameworks:** PyTorch, LangChain, FastAPI, Flutter, Unity
- **Databases:** PostgreSQL, MongoDB, ChromaDB, FAISS
- **Miscellaneous:** Docker, Prometheus, Git, Postman, Android Studio, LaTeX
- **Languages:** Bengali (Native), English (Advanced), Hindi (Verbal), German (A1.1)
- **Standardized Test Scores:** GRE - 307, IELTS - 7.5 (2024)

Selected Projects

Cold Email Generator for Graduate applications
Python, LangChain, FAISS, RAGAS, Llama-4 **April 2025**

- Automated and end-to-end multimodal RAG system to streamline graduate application outreach by aligning applicant profiles with professor research via web scraping, text parsing, and CV image analysis, achieving a **92.6% relevance score** (RAGAS).
- Designed a LangChain-based pipeline leveraging LLaMA-4 Maverick, Jina Embeddings v3, and FAISS, leveraging both textual and visual embeddings to generate contextually personalized cold emails.
- Incorporated Cohere Reranker v3.5, achieving a **Precision@K of 87.3%**, with robustness validated across diverse input formats (PDFs, HTML, images).

NoSmokeZone: AI for Smoker Detection in Public Spaces
Python, TensorFlow, Keras, OpenCV **December 2023 – January 2024**

- Engineered a real-time smoker detection system with minimal human intervention leveraging CNN models (EfficientNetV2, VGG16, ResNet-50) and Vision Transformer, achieving **93% accuracy**.
- Fine-tuned models and applied data augmentation techniques, reducing false positives/negatives by **around 7%** compared to baseline models, significantly minimizing misclassifications.

Malicious Website Detection Using Machine Learning
Python, Flask, HTML/CSS, JavaScript **October 2022 – November 2022**

- Implemented and deployed a machine-learning model achieving **94% accuracy** for real-time malicious website detection.
- Trained and benchmarked **4 models** — KNN, SVM, LR & MLP — and optimized hyperparameters, resulting **14% enhancement** in model performance .
- Built a Chrome extension using JavaScript to extract **27 real-time features** (e.g., URL length, domain authority) from webpages and integrated user feedback collection, used by **100+ test users**.
- Applied a continual learning pipeline that retrains the model every **24 hours** on the feedback data, reducing false positives by **almost 6%** and increasing detection robustness.

Open Source Contributions

- `Huggingface/transformers`

Achievements

2020 Recipient of the prestigious ICCR Scholarship by the Ministry of External Affairs, Govt.of India for academic excellence and promoting cultural exchange.
2017 SSC Board Merit Order Scholarship from the Chamber of Commerce, Kushtia, Bangladesh.
2015 JSC General Grade Scholarship from the Government of Bangladesh.
2012 1st position, Zilla Shilpakala Academy Kushtia Art Competition for Independence Day.
2011 2nd position, Bangladesh Udichi Shilpigoshthi Kushtia Art Competition for Bengali New Year.

References

Dr. M. Brindha, Associate Professor
Department of Computer Science & Engineering, NIT Trichy, India | brindham@nitt.edu

Amit Sethi, Professor
Department of Electrical Engineering, IIT Bombay, India | asethi@iitb.ac.in

Last updated in August 2025