# Arpon Kapuria

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

## Education

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

**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

## Research Interests

Generative NLP, Representation Learning, Reasoning, Factuality and Control

## Experience

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

- Research on multimodal language modeling with an emphasis on reasoning and trustworthiness.

**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 frameworks and 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**.
- Evaluated the CRAFT model for text detection in X-ray images and extracting clinically relevant annotations.

**National Institute of Technology Tiruchirappalli** Tiruchirappalli, India
Undergraduate R&D 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, Bash
- **Frameworks & Libraries:** PyTorch, LangChain, Hugging Face, FastAPI, Flutter, Unity
- **Databases & Storage:** PostgreSQL, MongoDB, Vector Databases
- **MLOps:** Docker, Kubernetes, MLflow, DVC, GitHub Actions, Prometheus, Grafana
- **Tools & Platforms:** Linux, Git, Postman, Android Studio, LaTeX
- **Languages:** Bengali (Native), English (IELTS 7.5), Hindi (Verbal), German (A1.1)

## Selected Projects

**Cold Email Generator**
Python, LangChain, FAISS, RAGAS, Llama-4, BeautifulSoup (BS4) **April 2025**

- Automated an 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 (Groq AI)**, **Jina Embeddings v3**, and **FAISS**, leveraging multimodal 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, FastAPI, Docker **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.
- Built API endpoints and deployed the system for reliable **model serving** using **FastAPI**, containerized with **Docker** for scalability and smooth integration.

**Malicious Website Detection Using Machine Learning**
Python, Scikit-Learn, Flask, Docker, JavaScript **October 2022 – November 2022**

- Implemented and deployed a machine-learning model via **Flask** and **Docker**, 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
- LangChain/langchain
- PyTorch/pytorch

## Achievements

**2020** **ICCR Scholarship** – Ministry of External Affairs, Government of India; awarded for academic excellence and funded my bachelor studies.
**2017** **Academic Excellence Award** – Chamber of Commerce, Kushtia; SSC Board (10th) Results.
**2015** **General Grade Scholarship** – Government of Bangladesh; JSC Board (8th) Results.
**2012** **1st Place** – Zilla Shilpakala Academy, Kushtia; Art Competition (Independence Day).
**2011** **2nd Place** – Bangladesh Udichi Shilpigoshthi, Kushtia; Art Competition (Bengali New Year 1418).

## 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