Arpon Kapuria

arpon-kapuria | arpkapuria@gmail.com | GitHub | LeetCode | +880 1797 288296

Summary

A recent graduate in Computer Science with a deep passion for technology, research, and teaching. My goal is to inspire the next generation of learners and augment human intelligence with more capable artificial intelligence to improve our day-to-day life. Outside of this, I like traveling.

Research Interests

Deep Learning, Representation Learning, Reasoning in AI Models (Language + Image)

Education

National Institute of Technology, Tiruchirappalli **December 2020 – June 2024**
B.Tech in Computer Science & Engineering | CGPA 7.75/10 (First Class) *Tamil Nadu, India*

Kushtia Government College, Kushtia **2019**
HSC (Science) | GPA 5/5 *Kushtia, Bangladesh*

Kushtia Zilla School, Kushtia **2017**
SSC (Science) | GPA 5/5 *Kushtia, Bangladesh*

Relevant Courseworks

- Operating Systems
- Computer Networks
- Database Management Systems
- Data Structures and Algorithms
- Technical Writing
- Linear Algebra and Calculus
- Machine Learning Techniques
- Augmented and Virtual Reality
- Artificial Intelligence
- Deep Learning Techniques
- Natural Language Processing
- Image Processing and Applications

Experience

Advanced Machine Intelligence Research Lab **February 2025 - Present**
Research Intern *Dhaka, Bangladesh*

- Investigating LLM reasoning capabilities to improve performance on complex tasks, with a focus on reducing hallucinations and achieving human-like decision-making.

Indian Institute of Technology Bombay **May 2023 – July 2023**
Research Intern | MeDAL Lab *Mumbai, India*

Project 1: Enhancing Self Supervised Learning framework - BYOL

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

Project 2: Radiology DICOM Image Anonymization

- Created an user-friendly Flask-based system for DICOM image processing and anonymization using Pydicom, ensuring medical data privacy.

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

Project: Centralized Power Cluster Home Automation

- Developed a Flutter app to automate the operations and billing for a power cluster in Assam, India, serving 200 homes.
- Integrated Flutter front-end with API developed in Django and used an MQTT server for IoT communication.
- This project was funded under SUSTENANCE, a Government of India initiative for carbon-neutral energy communities.

Projects

Cold Email Generator for Graduate applications **April 2025**

Python, LangChain, FAISS, Llama-4, Jina Embeddings v3, Cohere Reranker v3.5

- Automated a LangChain based Retrieval Augmented Generation (RAG) system that scrapes and processes data from professor and applicant websites, matching research interests for graduate applications.
- Designed a pipeline that leverages Llama-4 Maverick LLM, Jina embeddings and vector database (FAISS) to generate personalized and contextually relevant cold emails.
- Incorporated Cohere reranker to fine-tune similarity searches, boosting the precision and impact of the emails for improved response rates.

NoSmokeZone – AI for Smoker Detection in Public Spaces **December 2023 – January 2024**

Python, TensorFlow, Keras, OpenCV

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

Malicious Website Detection Using Machine Learning **September 2022 – November 2022**

Python, Flask, JavaScript, Chrome Extension tools

- Implemented a machine-learning model with around 94% accuracy to detect malicious websites.
- Trained and compared various classification algorithms, including supervised and neural network-based methods, to achieve a robust performance.
- Built a chrome extension using JavaScript to extract the features from the webpage to test the website against the trained model to classify if it is malicious or not. Added a feature to receive user feedback.
- Adopted a continual learning approach to store the user feedback and retrain the model daily to improve performance.

Skills

Programming Languages: C, C++, Python, Dart, JavaScript, SQL
Frameworks: PyTorch, LangChain, FastAPI, Flutter, Unity
Database: PostgreSQL, MongoDB, ChromaDB, FAISS
Miscellaneous: Linux, Docker, Prometheus, Git, VS Code, Android Studio, Google Colab, Latex
Languages: Bengali (Native), English (Proficient), Hindi (Verbal), German (A1.1)
Standarized Test Scores: GRE - 307, IELTS - 7.5 (2024)

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 organized Art competition for Independence Day.
2011 2nd position, Bangladesh Udichi Shilpigoshthi Kushtia organized Art competition for Bengali New Year ' 1418.

References

1. Dr. M. Brindha
   Associate Professor,
   Department of Computer Science & Engineering
   NIT Trichy, India - 620015
   `brindham@nitt.edu`

2. Amit Sethi, PhD
   Professor, Department of Electrical Engineering
   IIT Bombay, India - 400076
   `asethi@iitb.ac.in`