RAKIB AHSAN

iamrakib242@gmail.com rab2807.github.io rakib-ahsan-75b172309 rab2807

RESEARCH INTERESTS

Trustworthy AI | Adversarial Prompting | Computer Vision

WORK EXPERIENCE

Lecturer, Department of CSE, United International University

April 2025 – Present

Courses Teaching : Data Structures and Algorithms | Discrete Mathematics | Operating Systems | Operating Systems Lab | Computer Networks Lab

EDUCATION

Bachelor of Science in Computer Science & Engineering

Bangladesh University of Engineering & Technology

February 2020 – March 2025

CGPA : 3.90/4.00

Notable Courses : Computer Security | Machine Learning | Linear Algebra | Operating Systems | Compiler | Computer Networks | Artificial Intelligence | Database

RESEARCH EXPERIENCE

SequentialBreak: Large Language Models Can be Fooled by Embedding Jailbreak Prompts into Sequential Prompt Chains

Trustworthy AI | Large Language Models | 2024 - 2025

- With collaboration with Md Rafi Ur Rashid (Penn State University), we developed a novel automated jailbreak attack for LLMs. Our methodology exploited the *attention dilution* phenomenon by embedding the harmful prompt within a series of benign prompts in a single query. We showed three different attack scenarios where sequential prompt structure can be used to elicit harmful responses. While existing jailbreak methods often require multiple rounds for success, our one-shot attack achieved a substantially high attack success rate against recent LLM models.

Accepted in ACL Student Research Workshop (SRW) 2025 [DOI] [Slides]

A Survey of Attacks and Defenses in LLM Agents and Vision-Language Models

Trustworthy AI | Vision Language Models | Agentic AI | 2025 - Ongoing

- Ongoing project on a comprehensive survey in collaboration with Md Rafi Ur Rashid (Penn State University), examining the landscape of attacks and defenses targeting large language model (LLM) agents and vision-language models. The survey covers a broad spectrum of security threats, including adversarial attacks, jailbreaks, data poisoning, backdoors, privacy breaches, and model extraction attacks.

AttMetNet: Attention-Enhanced Deep Neural Network for Methane Plume Detection in Sentinel-2 Satellite Imagery

Computer Vision | Remote Sensing | 2024 - 2025

- This is my undergraduate thesis project under Dr. Tanzima Hashem. Methane is the second-largest contributor to greenhouse gas emissions, and its reduction offers a rapid path to limiting global temperature rise. This work focuses on generating segmentation masks of methane plumes from multispectral satellite imagery which is challenging due to limited data and complex plume morphology. We enhanced a U-Net architecture with attention gates and incorporated spectral indices which outperforms the state-of-the-art detection methods in real-world conditions.

Under preparation for "NeurIPS 2025 Workshop: Tackling Climate Change with Machine Learning"

RELEVANT COMPETITIONS

IEEE SPS Signal Processing Cup 2025 | Site Link

First Runner-up [arxiv]

- Our methodology employs MaxViT, CoAtNet, and EVA-02 as backbones fine-tuned using supervised contrastive loss to enhance feature separation. A majority voting ensemble is done on the classification heads based on backbones to get the final prediction. The supervised contrastive loss creates well-separated embedding spaces, which significantly enhances the classification process.

NOTABLE ACADEMIC PROJECTS

Coastline Segmentation with Satellite Imagery | [github] Pytorch SentinelHub OpenCV Earth Engine

- A deep learning semantic segmentation task to differentiate coastline land and water area. 12-channel Sentinel-2 images were pre-processed and fed to different segmentation models for comparison. Unet, Unet with attention and transformer-based segmentation models were becnhmarked.

Deep Neural Network from Scratch | [github] Numpy Seaborn Matplotlib

- A complete neural network framework implemented from scratch using only NumPy without using high-level frameworks like TensorFlow or PyTorch. Implemented forward/backward propagation, activation functions (ReLU, Sigmoid, SoftMax), loss functions (Cross-entropy loss), Optimizers (SGD and Adam), dropout, batch normalization and early stopping.

DormEase: A Dormitory Management System | [github] NodeJS ReactJS Tailwind CSS Supabase

- DormEase is a dormitory management system simplifying the processes of room allotment, mess managing and complaint resolution. Residents and administrators can communicate with a dedicated new feed and noticeboard.

TCP Adaptive Reno with NS-3 | [github] NS-3

- With NS-3, TCP Adaptive Reno algorithm is implemented that adaptively adjusts its behavior based on network conditions. TCP Adaptive Reno extends TCP Westwood+ with RTT-based congestion estimation and dynamic window sizing. A dumbbell topology was used for comparing TCP algorithms.

TECHNICAL SKILLS

- **Programming Languages:** C/C++, C#, Java, Python, Javascript, x86 Assembly, SQL, Bash
- **Machine Learning** Pytorch, Numpy, Pandas, Gymnasium, Matplotlib, Scikit-learn, Seaborn, OpenCV
- **Tools & Softwares:** Git, Shell, LaTeX, Cisco Packet Tracer, Unity Engine, WireShark, Metasploit
- **Frameworks:** NodeJS, ReactJS, Bootstrap, Tailwind CSS, P5.js

ACADEMIC ACHIEVEMENTS

Dean's List Scholarship

- Recipient of Dean's List Scholarship in every semester for academic excellence.

Merit Scholarship

- Recipient of Merit Scholarship in January 2021, July 2021, January 2022 and July 2023 sessions for achieving top academic result.

LEADERSHIP

BUET CSE Fest 2024 Game Jam

Organizer

- Co-organized the Game Jam event.
- Took an workshop on Game Development using Unity for the same event.