My research interests lie at the intersection of Efficient Deep Learning, Trustworthy AI, and Secure Machine Learning. I intend to focus on improving and protecting multimodal large language models (LLMs) and distributed learning systems. My goal is to create AI frameworks that are efficient, privacy-preserving, and reliable so they can be safely used in real-world and resource-limited environments.

As deep learning models grow larger and more multimodal, they face major challenges in efficiency, privacy, and trust. Training or fine-tuning these models often requires sensitive or private data, which makes them open to risks like data leakage, gradient inversion, or membership inference attacks. At the same time, their high computational cost limits practical and large-scale deployment. I want to solve these problems by designing methods that allow efficient, private, and safe training, adaptation, and unlearning in large AI systems.

I am motivated by some key questions: How can we build multimodal systems that can reason over different types of data for specific tasks, while ensuring strong privacy and security? And how can we make sure these models are not only accurate but also robust against attacks and able to unlearn data securely when needed? Through my research, I try to answer these questions by developing practical solutions that connect advanced AI techniques with real-world applications.

Interdisciplinary Research Background

My research journey started with developing a multimodal medical assistant for dermatology under the supervision of Dr. Mohammad Ariful Haque, Professor at BUET. At first, I fine-tuned a vision-language model on the Dermnet dataset and built a chatbot that performed reasonably well. However, I noticed a problem, the model gave fluent answers that were not always clinically accurate, showing poor structured reasoning. I realized that simple fine-tuning was not enough.

To solve this issue, I used Grouped Relative Policy Optimization (GRPO) and Direct Preference Optimization (DPO) to improve the model's reasoning process. These methods helped the model generate more structured and diagnostic responses, instead of just fluent ones. As a result, the chatbot became better at both patient-like interaction and diagnostic accuracy. To further improve its performance, I added DINOv2 for feature extraction and a knowledge graph-based retrieval system (RAG). Since medical systems often face limited resources and privacy concerns, I also applied structured pruning to make the model lighter and more suitable for real-world use.

This work led to our submissions of CLARIFY [1], a specialist-generalist framework; GRPO++ [2], which studies reasoning in low-resource conditions; and a paper on Compression Strategies[3] for Efficient Multimodal LLMs in Medical Contexts. These works reflect my focus on creating practical, efficient, and privacy-aware AI systems. This project taught me the full process of AI research, from identifying a key problem to developing real, scalable, and secure solutions.

My commitment to building trustworthy AI is further demonstrated by my accepted papers on a transfer learning approach for skin cancer classification from imbalanced data [4] and a multi-stage deep learning method for tuberculosis detection with explainable insights [5]. These projects underscore my core belief that a model's prediction is only as valuable as the clinician's ability to trust and interpret it, and that this trust is inextricably linked to data security and privacy.

AI in Practice and Leadership

As a Machine Learning Engineer at Advanced Chemical Industries Ltd., I have turned research ideas into practical applications by developing and deploying robust systems. These include an LLM-powered CV Sorter, a tabular data analysis tool called Insight Explorer, and a computer vision system for fraud detection. These experiences have taught me how to build scalable and maintainable systems, skills that are directly useful for creating secure and distributed machine learning frameworks.

Beyond my technical work, I have also taken on leadership roles. As the first-ever Student Executive for ACM SIGCOMM, I worked with the Chair, Dr. Matthew Caesar (Professor, CS, UIUC), to lead initiatives for the global networking community. My contributions included helping develop the official SIGCOMM website and co-founding a paper reading group.

Where I See Myself

My experiences have shaped my goal to build AI systems that are not only intelligent and interpretable but also secure, privacy-preserving, efficient, and robust. After completing my Ph.D., I hope to pursue a career in academia as a professor, leading a research group focused on developing trustworthy and privacy-centered AI for diverse areas.

A Ph.D. at UC Riverside, under the guidance of Prof. Guler, would give me the right environment to gain deeper knowledge and research skills to achieve this goal. I believe my background in both theoretical research and practical model deployment will help me make meaningful contributions to your department. I am especially motivated by the chance to create AI solutions that connect machine intelligence with human well-being, while protecting data privacy and security.

References

[1] **Aranya Saha***, Tanvir Ahmed Khan*, Ismam Nur Swapnil*, Mohammad Ariful Haque, “CLARIFY: A Specialist-Generalist Framework for Accurate and Lightweight Dermatological Visual Question Answering,” Under Review at *IEEE Transactions on Human-Machine Systems*. [arXiv:2508.18430]

[2] Ismam Nur Swapnil*, Tanvir Ahmed Khan*, **Aranya Saha***, Mohammad Ariful Haque, “GRPO++: Enhancing Dermatological Reasoning Under Low-Resource Settings,” Under Review at *IEEE Journal of Biomedical and Health Informatics*. [arXiv:2510.01236]

[3] Tanvir Ahmed Khan, **Aranya Saha**, Ismam Nur Swapnil, Mohammad Ariful Haque, “Compression Strategies for Efficient Multimodal LLMs in Medical Contexts,” Under Review at *Springer Journal of Signal Processing Systems*. [arXiv:2507.21976]

[4] Shadman Sobhan, **Aranya Saha**, Tanvir Ahmed Khan, Abduz Zami, “Skin Cancer Classification Using Pre-trained CNNs: A Transfer Learning Approach Addressing Imbalanced Data Challenges,” in *Proc. 2nd Int. Conf. on Next-Generation Computing, IoT and Machine Learning (NCIM)*, June 2025. [IEEE Xplore]

[5] Shadman Sobhan, Abduz Zami, Mohiuddin Ahmed, Tanvir Mahtab Zihan, Tanvir Ahmed Khan, **Aranya Saha**, “A Multi-Stage Deep Learning Approach to Tuberculosis Detection with Explainable Insights,” in *Proc. 2nd Int. Conf. on Next-Generation Computing, IoT and Machine Learning (NCIM)*, June 2025. [IEEE Xplore]

*denotes co-first authorship.