My primary research interest lies in software engineering, more specifically in empirical software engineering and software security. I have recently graduated from Bangladesh University of Engineering and Technology (BUET) with a Bachelor of Science in Computer Science and Engineering. My undergrad thesis focused on removing software vulnerabilities in source codes, in which I created a dataset to use a novel approach to refactor code containing security vulnerabilities. This research project helped me to understand the broadness of research in software engineering. It motivated me to study further on this topic and explore the different perspectives of this field through a Ph.D. program. I am especially interested in software vulnerabilities, code refactoring, and machine learning for software engineering. Nevertheless, I am open to any new ideas.

During my undergrad, I was supervised by Dr. Anindya Iqbal to study SQL injection vulnerabilities. This problem was suggested by a US-based software startup company working on intelligent code repairing tools. This study aimed to automatically repair SQL injection vulnerabilities that originated from the unparameterized query in the source code. We introduced a learning-based approach previously described by researchers from Facebook. In this approach, we used hierarchical clustering to cluster similar vulnerable codes and their fixes. Later, our test set used this cluster to find possible spots to remove SQL Injection vulnerabilities. My role in this study was to mine open-source codes from GitHub repositories to create a dataset containing vulnerable Java codes and their fixes. We converted the code into an Abstract Syntax Tree (AST), which helped generate a language agnostic solution. My thesis partner and I reproduced the code from the motivating paper of Facebook researchers and applied it to the dataset. Later, I used the approach in PHP codes to establish the idea of language independence of our system. We compared the result with the previous rule-based approach. The main contribution of this study was to apply a novel approach to solve another common software vulnerability and produce a language-agnostic tool that could suggest fixes from a learning-based technique. The challenging parts of this project are reproducing code from the state-of-the-art paper, extending our approach to a generic tool though it needs language-dependent preprocessing. This research has been accepted to SANER, 2021, one of the top conferences in the Software Engineering domain.

Recently, I was working under Dr. Gias Uddin from the University of Calgary. In this study, I examined common security flaws in the IoT domain. I cloned GitHub repositories with "IoT" tags and applied basic rule-based techniques to filter out potentially vulnerable source codes. Then we created a benchmark dataset to improve our tool by investigating the Abstract Syntax Tree (AST) level and applying shallow machine learning models. Besides that, I guided one of his undergrad research interns to study overflow-related software vulnerabilities in the IoT domain. I managed meetings with him to provide insight into my approach and write skeleton codes to extend his ideas with a similar codebase.

Though my research interest is in software engineering, I have explored other domains of computer science. I worked with Dr. Shubhra Kanti Karmaker from Auburn University to create question clusters from asked questions in MOOCs. This study dealt with big data, applied machine learning, and document clustering. This research would help me use a machine learning approach to big data, like source codes, project metadata, etc. Besides that, I worked on finding the challenges and opportunities of designing telemedicine

solutions from the perspective of a developing country under the supervision of Dr. Alim Al Islam, BUET. In this project, our team created a telemedicine solution by following a user-centric design pipeline. We analyzed existing solutions in the market and surveyed the users to know the situation of telemedicine solutions in Bangladesh. The findings of this study have been submitted to the ACM CHI Conference on Human Factors in Computing Systems. This study could help me set up questioners for real-life software users and design empirical software engineering studies. My research experiences in Big Data, Applied Machine Learning, and Human-Computer Interaction motivate me to interconnect different computer science domains and provide an in-depth understanding of different research viewpoints.

These experiences motivated me to study software security in source code, helped me handle big data like source code mining, create language-agnostic code repairing tools, and the importance of human factors in the research study. In the graduate study, I expect to work on a broad area of software engineering, not limited to these areas, but also architectural software flaws in design decisions and system engineering problems.

I love to enjoy my understanding of programming languages and software engineering methods in research projects. I am always open to exploring the various aspects of software engineering and their impacts on industrial software engineering practice. My experience helped me to feel that research work should have a real-life application. My thesis proposal came from the industry. I mostly worked on the open-source software code base, security vulnerabilities, applied machine learning, and human factors in design philosophy, impacting the industry. Through my Ph.D. study, I expect to become an excellent researcher who makes a difference in industry and academia under proper guidance and hard work.