

# Mushtari Sadia

[LinkedIn](#) | [GitHub](#) | Email: [mushtarisadia98@gmail.com](mailto:mushtarisadia98@gmail.com)

**Research Interests:** Computer Security, Privacy, Machine Learning

Location: Dhaka, Bangladesh

Phone: +8801746518030

## WORK EXPERIENCE

### Lecturer

Department of Computer Science and Engineering, BRAC University

Department of Computer Science and Engineering, BUET (Adjunct Lecturer)

Dhaka, Bangladesh

Jun 2023-Current

Nov 2023-Mar 2024

### Research Fellow (Part-time)

Supervisor: Dr. Praneeth Vepakomma, MIT Media Lab (Fatima Fellowship)

Remote

Aug 2023-Current

## EDUCATION

### Bangladesh University of Engineering and Technology (BUET)

B.Sc. in Computer Science and Engineering; CGPA: 3.84/4.0

Dhaka, Bangladesh

April 2018 - May 2023

#### Notable Courses:

Computer Security, Machine Learning, Computer Architecture, Microprocessors, Microcontrollers & Embedded Systems, Computer Networks, High Performance Database Systems, Operating Systems, Computer Graphics, Compiler Design, Introduction to Bioinformatics

## TECHNICAL SKILLS

**Languages** : C/C++, x86 Assembly, Bison/Flex, Python, Java, Javascript, Bash, MySQL

**Frameworks** : Docker, PyTorch, NS3, xv6, Django REST, ReactJS, Git, Oracle DBMS, LaTeX, Wireshark

**Libraries** : Sklearn, Pandas, Matplotlib, Seaborn

## COMPETITIONS & ACHIEVEMENTS

### Microsoft Virtual Hackathon

1st Runner Up

2022

[Github](#)

- I co-created an AI based dengue forecast system using **MS Azure Services** that predicts the number of dengue cases in any given region based on the recent cases in that region and the state of different weather parameters.

### HerWILL Datathon

Champion

2022

[Github](#)

- A machine learning based forecasting system for predicting taxi demand in a city.

### UNDP Women's Digital Innovation Hackathon

2nd Runner Up

2021

- Built an AI based Dengue Monitor & Control System

**Also Participated In:** Dhaka-AI 2020, Ada Lovelace Datathon 2021, NLP Hackathon 2023, Robi Datathon 2.0

## HONORS & AWARDS

- Fatima Fellowship:** The Fatima Predoctoral Fellowship is a 9-month program in which students who are planning on applying to CS, PhD programs in the US work with researchers. I'm currently working on a project of the [MIT Camera Culture Group](#) through this fellowship.
- Microsoft for Startups Founders Hub Membership:** My AI-based dengue forecast system prototype built with MS Azure received a 3000 USD startup funding along with a membership in MS Startups Founders Hub.
- GHC 22 Scholarship:** A scholarship granted to only a few selected candidates based on merit. The scholarship covers tuition and other costs related to attending the Grace Hopper Celebration of Women in Computing, a conference that brings together women in tech from around the world.
- Dean's List Scholarship (2018-2021):** Scholarship granted to undergraduate students for their academic excellence.
- Dhaka Board General Scholarship (HSC)** Scholarship granted for outstanding performance in the secondary education qualification exam of Bangladesh.

## RESEARCH EXPERIENCE

---

### Effectiveness of Language Models in Detecting Advanced Persistent Threats from Provenance Graphs (2022 - Current)

Computer Security, Natural Language Processing

[\[arXiv link\]](#)

Undergraduate thesis project in collaboration with National Security Council, Canada. In this work, I co-implemented a framework which includes a robust process of creating a provenance graph from system log data, subsequently generating event sequences from that graph and finally transforming the data into a suitable format for transformer-based language models. My personal contributions were building the postgres database from the raw log datasets, designing SQL queries to generate the provenance graph, co-writing the code for preprocessing the graph data to extract traces with relevant attributes, and finally building the experiments with various pre-trained LLMs. We were able to achieve state of the art performance from our framework in detection of APT attacks. Currently, this work is under review for publication.

### Advancing Parallelization In Deep Learning Training: A Novel HSIC-Based Approach (2023 - Current)

Federated Learning, Privacy, Optimization

Research project in collaboration with the MIT Camera Culture Group. I developed a method for parallelizing the forward propagation in neural network training, utilizing the HSIC objective function to eliminate the need for backpropagation.

### Development of Flood Forecasting System for Bangladesh-India Using Different Machine Learning Techniques (2020)

Machine Learning

In this study, five different machine learning algorithms were used to develop total 180 independent models based on a different combination of time lags for input data and lead time in forecast. Models were developed for Someshwari-Kangsa sub-watershed of Bangladesh's North Central hydrological region with 5772  $km^2$  drainage area. I was responsible for preprocessing the datasets and applying the ML models.

- Poster Presentation: Haque, M. H., **Sadia, M.**, and Mustaq, M.: EGU General Assembly 2021, online, 19–30 Apr 2021, EGU21-15294, <https://doi.org/10.5194/egusphere-egu21-15294>, 2021.

### AI based Dengue Forecasting System (2020-2022)

[Github](#)

Machine Learning

In this project, I co-created a time series forecasting model-based dengue forecast system that predicts the number of dengue cases in any given region based on the recent cases in that region and the state of different weather parameters.

We presented the idea and won the **1st Runner Up** position in **Microsoft Virtual Hackathon 2022**.

## ACADEMIC PROJECTS

---

### Privacy Preserving ML with Secure-Sum Protocol

Docker, Numpy

[GitHub](#)

- Implementation of a secure-sum gradient descent protocol for Federated Learning using Docker

### Forge, A Complete Research Tool

React.js, Django REST, PostgresDB

[YouTube](#)

- Designed a system by creating BPMN, ERD, Collaboration, Sequence diagrams, then according to the design, built a research-management website, Forge, with Django-REST framework, Docker, PostgresDB, React-JS. Forge streamlines research management by providing a comprehensive solution for exploring, collaborating, and submitting research projects.

### TcpSwift, A Congestion Control Algorithm For Datacenters

NS3

[GitHub](#)

- A comparative study on the varied network parameters and performance metrics of two built-in and one implemented novel congestion control algorithms, with three different network topologies simulated with **NS3**.

## LEADERSHIP EXPERIENCE

---

### President, Bangladeshi Women In Computer Science & Engineering (BWCSE)

April 2022 - May 2023

The mission of BWCSE is to empower Bangladeshi women by fostering academic, social, and professional growth in the field of computer science and engineering. As president, my responsibilities included coordinating and organizing various competitive programming competitions, workshops and seminars on various CS fields.

### Batch Representative, Bangladeshi Women In Computer Science & Engineering (BWCSE)

Apr. 2021 - Apr. 2022

## CERTIFICATIONS

---

Coursera: [Convolutional Neural Networks](#), [Improving Deep Neural Networks: Hyperparameter tuning, Regularization and Optimization](#), [Structuring Machine Learning Projects](#), [Neural Networks and Deep Learning](#)