Mohammed Latif Siddiq

150C Fitzpatrick Hall of Engineering, University of Notre Dame, Notre Dame, IN 46556

+1 (813) 331-7991 | lsiddiqsunny@gmail.com | https://lsiddiqsunny.github.io

GitHub | LinkedIn | Google Scholar | Researchgate

EXECUTIVE SUMMERY

- Software engineering researcher with two years of hands-on experience developing and conducting research in the intersection of software security and applied machine learning.

- Industry-level software engineering experience developing enterprise and machine learning-based software.

EDUCATION

University of Notre Dame   January 2022 - Present
Ph.D. in Computer Science and Engineering
Department of Computer Science and Engineering   Running CGPA: 3.722/4.00

Bangladesh University of Engineering and Technology   February 2016 - February 2021
Bachelor of Science in Computer Science and Engineering
Department of Computer Science and Engineering   CGPA: 3.46/4.00

PROFESSIONAL EXPERIENCES

University of Notre Dame, IN, USA | *Graduate Research Assistant*   January, 2023 - Present

- Working in the intersection of language models, their application, and security.
- Published four peer-reviewed research papers.

Cummins Inc., IN, USA | *Data Scientist Co-op*   June, 2023 - August, 2023

- Worked on building a question-answering pipeline using large language models.
- **Tech stack**: Python, Langchain, Gradio, and Databricks platform.

University of Notre Dame, IN, USA | *Graduate Teaching Assistant*   January, 2022 - December, 2022

- **Spring 2022**: CSE 30321 Computer Architecture
- **Fall 2022**: CSE 30332 Programming Paradigms

REVE Systems, Dhaka, Bangladesh | *Junior Software Engineer*   March, 2021 - November, 2021

- Developing enterprise software for Bangladesh Parliament Secretariat.
- **Tech stack**: Java servlet, JSP and MySQL.

PEER REVIEWED PUBLICATIONS

22nd **IEEE International Working Conference on Source Code Analysis and Manipulation (SCAM)**   2022
An Empirical Study of Code Smells in Transformer-based Code Generation Techniques, **Mohammed Latif Siddiq**, Shafayat Hossain Majumder, Maisha Rahman Mim, Sourov Jajodia, and Joanna C. S. Santos

28th**IEEE International Conference on Software Analysis, Evolution and Re-engineering (SANER)**   2021
SQLIFIX: Learning Based Approach to Fix SQL Injection Vulnerabilities in Source Code, **Mohammed Latif Siddiq***, Md. Rezwanur Rahman Jahin*, Mohammad Rafid Ul Islam, Rifat Shahriyar, and Anindya Iqbal
(*Equal contribution)

WORKSHOP PUBLICATIONS

2nd Intl. Workshop on Natural Language-based Software Engineering 2023
Zero-shot Prompting for Code Complexity Prediction Using GitHub Copilot , **Mohammed Latif Siddiq**, Abdus Samee, Sk Ruhul Azgor, Md. Asif Haider, Shehabul Islam Sawraz, and Joanna C. S. Santos

1st Intl. Workshop on Mining Software Repositories Applications for Privacy and Security 2022
SecurityEval Dataset: Mining Vulnerability Examples to Evaluate Machine Learning-Based Code Generation Techniques, **Mohammed Latif Siddiq**, and Joanna C. S. Santos

1st Intl. Workshop on Natural Language-based Software Engineering 2022
BERT-Based GitHub Issue Report Classification, **Mohammed Latif Siddiq**, and Joanna C. S. Santos

OTHER PUBLICATIONS

ArXiv Preprint 2023
A Lightweight Framework for High-Quality Code Generation(Preprint), **Mohammed Latif Siddiq**, Beatrice Casey, Joanna C. S. Santos

ArXiv Preprint 2023
Exploring the Effectiveness of Large Language Models in Generating Unit Tests(Preprint), **Mohammed Latif Siddiq**, Joanna C. S. Santos, Ridwanul Hasan Tanvir, Noshin Ulfat, Fahmid Al Rifat, and Vinicius Carvalho Lopes

ArXiv Preprint 2022
Shashthosheba: Dissecting Perception of Bangladeshi People towards Telemedicine Apps through the Lens of Features of the Apps(Preprint), Waqar Hassan Khan, Md Al Imran, Ahmed Nafis Fuad, **Mohammed Latif Siddiq**, and ABM Islam

IEEE Computer Society Bangladesh Chapter Winter Symposium 2020
Bangla Captioning Image Taken by Blind People(Extended Abstract), **Mohammed Latif Siddiq**, and Nafis Tahmid Chowdhury

Bangladesh Blockchain Olympiad 2022
Localized Sustainable and Ecofriendly Energy Generation and Distribution Using Blockchain Network: Bangladesh Perspective(Poster), Md. Abdullah Mia, Rashik Ahnaf, **Mohammed Latif Siddiq**, and Md. Mahmudur Rahman Sayem

4th International Conference on Networking, Systems, and Security (NSysS) 2017
Online Blood Bank-Connecting Donors and Blood Needing People in Bangladesh(Poster), **Mohammed Latif Siddiq**, Aditya Chakma, and Farhan Tanvir Uthshaw

RESEARCH PROJECTS

Case Study: Wireshark
Open Source Project, Threat Modeling, CVE March, 2022

- Analyzed the architectural structure and threat modeling of an open-source packet capturing and visualizing tool, Wireshark.
- Decomposed the application, identified the threats, and documented a list of threats.

Telemedicine Sectors in Bangladesh
User-Centric Design, Human-Computer Interaction, Survey September, 2021

- Created a telemedicine solution for the project purpose
- Analyzed the top telemedicine solutions in the market and compared them with our solution

Detecting Security Bugs in GitHub Codes from the IoT Domain
Python, Deep Learning, GHTorrent July, 2021

- Worked towards uncovering security flaws in IoT-related GitHub code under Dr. Gias Uddin, Assistant Professor, University of Calgary.

Online Text Clustering for MOOCs | *Python, Deep Learning, Online Algorithm, Django, Nuxt* April, 2021

- Worked to establish a way to accomplish online text clustering for MOOC questions with Dr. Shubhra Kanti Karmaker, Assistant Professor at Auburn University in Alabama.

Bangla Image Captioning Captured by Blind People | *Keras, TensorFlow, Python, AWS* December, 2020

- Generated Bangla captions for the image captured by blind persons, formally known as the VizWiz dataset, using the deep learning model.
- One of the top 30 AI initiatives for the Bangladesh Government's AI for Bangla Competition.

Detecting Number of TCP and UDP Flows in SDN by ML
Machine Learning, Networking SDN, OpenFlow Protocol, Mininet, IPERF, Weka November 2020

- Developed machine learning-based methods to detect the number of TCP and UDP flows in a Software-Defined Network simulated in Mininet using IPERF to generate flows.

ACHIEVEMENTS

AI FOR BANGLA, BANGLADESH, 2021
Position: Top 30

GOOGLE HASHCODE, GLOBAL, 2019
Position: 5th Among Bangladeshi Teams

IEEEXTREME 12.0, GLOBAL, 2018
Position: 4th Among Bangladeshi Teams

AUB PROGRAMMING CONTEST, BANGLADESH, 2018
Position: 11th

SUMSUNG CODING CONTEST, BANGLADESH, 2018, 2019
Finalist

TECH FOR PEACE HACKATHON, BANGLADESH, 2017
Winner

SCHOLARSHIPS

TRAVEL GRANTS, ACM SIGSOFT TRAVEL GRANTS, 2023
1,200$ grants to participate in ICSE 2023 in-person and Student Mentoring Workshop in Melbourne, Australia.

TRAVEL GRANTS, BANGLADESH SWEDEN TRUST FUND, 2022-23
Around 150$ grants for one-way plane fare to attend USA school with financial funding.

TRAVEL GRANTS, NSF ICSE 2022 TRAVEL AWARDS, 2022
1,075$ grants to participate in ICSE 2022 in-person and Student Mentoring Workshop along with complimentary registration for the conference.

MELCHOR FELLOWSHIP, UNIVERSITY OF NOTRE DAME, 2022-27
Tuition scholarship including health insurance, payment of the technology, health center access fees, and bimonthly stipends.

TECHNICAL SCHOLARSHIP, BANGLADESH, 2016-20
Complimentary scholarship for regular engineering students.

GOVERNMENT SCHOLARSHIP, BANGLADESH, 2016-20
Awarded For Outstanding Performance in Higher School Certificate Examination.

GOVERNMENT SCHOLARSHIP, BANGLADESH, 2011-12
Awarded For Outstanding Performance in Junior School Certificate Examination.

TECHNICAL SKILLS

Languages: Java, C, C++, Python, Matlab, Dart, Shell
Database: Oracle, MySQL, SQLite, MongoDB, Firebase
OS: Windows, Ubuntu 20.04, Windows Server 2016, Mac OS Monterey, CentOS 7/RedHat
Version Control: Git(GitHub, Bitbucket, Gitlab), TFS(Azure DevOps)
Frameworks: Vue.Js, JSP, Flutter, JavaFX, Node.js, Django, Java servlet
Web Technology: HTML, CSS, Rest API, JSON, XML
Cloud: EC2, S3 Bucket, Azure
Technical Writing: LaTeX, Beamer, Overleaf
Other: Gradio, Langchain, Databricks, Google Apps Script, Software Defined Networking, PyTorch, OpenGL, Weka, Mininet, NS2

CERTIFICATES

TOEFL iBT
Total: 99(Out of 120, 30 Per section)
Reading: 24, Listening: 27, Speaking: 23, Writing: 25

LinkedIn Assessment
C, C++, Python, Django, MongoDB, MySQL, Git, HTML, Windows server, and Machine Learning

Other Certifications
Deep Learning Specialization(Coursera), Human-Computer Interaction I(Edx.org), Cyber Security Essential(CISCO Learning Platform), Problem Solving (Advanced) Skills Certification Test(Hackerrank)

Problem Solving
Codeforces(Highest Rating: 1620), Codechef(Highest Rating: 1841), Hackerrank(Highest Rating: 1780)

HIGHLIGHTED ACADEMIC COURSES

Graduate Course, University of Notre Dame
CSE-60770 Secure Software Engineering
CSE 60625 Advanced Machine Learning
CSE 60647 Data Science
CSE-60326 Computational Behavior Modeling
Undergraduate Course, Bangladesh University of Engineering and Technology
CSE-405 Computer Security
CSE-423 Fault Tolerant Systems
CSE-463 Introduction to Bioinformatics
CSE-471 Machine Learning
CSE-473 Pattern Recognition

REFERENCES

Dr. Joanna Cecilia da Silva Santos
Assistant Professor
Department of Computer Science and Engineering
University of Notre Dame, IN, USA.
Email: joannacss AT nd DOT edu
Relation: Ph.D. Advisor