

# M SAIFUL BARI

(+65) 85052376 ◇ bari0001@e.ntu.edu.sg

Ph.D. Student, Nanyang Technological University

sbmaruf.github.io

## EDUCATION

---

### Nanyang Technological University

Doctor of Philosophy (Ph.D.)

Natural Language Processing

Computer Science and Engineering

2019 - Present

CGPA: 4.63/5.0

### Islamic University of Technology

Bachelor of Science (BSc.)

Computer Science and Engineering

2012-2016

CGPA: 3.65/4.0

## SKILLS

---

### Computer Languages

Python, C/C++

### Tools & Library

Pytorch, Bash, Linux

### Problem Solving

Deep Learning, Algorithms & Data Structure

## HONORS & AWARDS

---

### Scholarship

NTU Research Scholarship, Fully funded Ph.D. scholarship for 4 years.

2019

### Scholarship

OIC Scholarship for undergraduate study, Islamic University of Technology

2012

### Champion

IUT Computer Programming Contest

2014

### Honorable Mention

Human Expedition on Mars Timeline 2018

2014

### Champion

IUT Computer Programming Contest

2015

### 2<sup>nd</sup>/100 in Inter University Programming Contest

Daffodill International University ACM ICPC world finals warmup contest 2016

2016

### 6<sup>th</sup>/100+ in Inter University Programming Contest

NSU Cybernavts National Programming Contest

2016

## EXPERIENCE

---

### Applied Scientist Intern (Part Time)

Amazon Web Service

July 2021 - Jan 2022

- Working on prompt tuning for multi-lingual models.

### Applied Scientist Intern

Amazon Web Service

Aug 2020 - Oct 2020

- Collaborated with Amazon Lex team.
- Worked on Cross-lingual Few-shot Adaptation.

## Research Assistant

Sep 2017 - Aug 2018

*Nanyang Technological University*

- Research on MT, NER and Adversarial Training.
- Maintain internal GPU servers
- Maintain group web-site

## Software Engineering Intern

Sep 2018 - Jan 2019

*Aubichol IT Limited*

- Worked on an early startup where the key responsibility was to design the architecture of sports analytic and Translation System.

## Lecturer

June 2017 - Sep 2017

*Southeast University*

- Teaching undergraduate CS student. Subject taught: \* Introduction to C \* Algorithms

## Software Engineering Intern

Nov 2015 - Dec 2015

*XeonBD*

- The course of the internship goes through Cloud Computing and kernel virtualization.

## Competitive Programmer

Nov 2012 - Nov 2016

*Islamic University of Technology*

- Solved different kinds of algorithmic and data structure based problem. Solve problem through codeforces, topcoder, lightoj, uva, codechef, hackerrank, projecteuler, coci, usaco. Solved around **2000 problems** in various online judge.

## Competitive Programming Trainer

June 2014 - Nov 2016

- Taught Competitive Programming in different Universities in Bangladesh.
  - Islamic University of Technology
  - Khulna University of Engineering and Technology
  - East West University
  - Dhaka University of Engineering and Technology

## PUBLICATION

---

1. Victor Sanh, Albert Webson, Colin Raffel, Stephen Bach, Lintang Sutawika, Zaid Alyafeai, Antoine Chaffin, Arnaud Stiegler, Arun Raja, Manan Dey, **M Saiful Bari**, Canwen Xu, Urmish Thakker, Shanya Sharma Sharma, Eliza Szczechla, Taewoon Kim, Gunjan Chhablani, Nihal Nayak, Debajyoti Datta, Jonathan Chang, Mike Tian-Jian Jiang, Han Wang, Matteo Manica, Sheng Shen, Zheng Xin Yong, Harshit Pandey, Rachel Bawden, Thomas Wang, Trishala Neeraj, Jos Rozen, Abheesht Sharma, Andrea Santilli, Thibault Fevry, Jason Alan Fries, Ryan Teehan, Teven Le Scao, Stella Biderman, Leo Gao, Thomas Wolf, and Alexander M Rush. **Multitask Prompted Training Enables Zero-Shot Task Generalization**. In *International Conference on Learning Representations, ICLR, 2022*
2. **M Saiful Bari**, Batool Haider, and Saab Mansour. **Nearest Neighbour Few-Shot Learning for Cross-lingual Classification**, 2021
3. **Bari, M Saiful**, Tasnim Mohiuddin, and Shafiq Joty. **UXLA: A Robust Unsupervised Data Augmentation Framework for Zero-Resource Cross-Lingual NLP**. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, ACL (Volume 1: Long Papers)*, pages 1978–1992, Online, August 2021. Association for Computational Linguistics
4. Tasnim Mohiuddin, **M Saiful Bari**, and Shafiq Joty. **AugVic: Exploiting BiText Vicinity for Low-Resource NMT**. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, Online, 2021. Association for Computational Linguistics

5. Tasnim Mohiuddin, **M. Saiful Bari**, and Shafiq R. Joty. **LNMap: Departures from Isomorphic Assumption in Bilingual Lexicon Induction Through Non-Linear Mapping in Latent Space**. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*, Virtual, November 2020
6. **M Saiful Bari**, Shafiq Joty, and Prathyusha Jwalapuram. **Zero-Resource Cross-Lingual Named Entity Recognition**. In *Proceedings of the 34th AAAI Conference on Artificial Intelligence*, AAAI '20, New York, USA, 2020. AAAI
7. Xiang Lin, Shafiq Joty, Prathyusha Jwalapuram, and **M Saiful Bari**. **A Unified Linear-Time Framework for Sentence-Level Discourse Parsing**. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, ACL '19, Florence, Italy, 2019. ACL
8. **M Saiful Bari**. *Regression for Data Analytics*, chapter 2, pages 33–54. CRC Press, Boca Raton, September 2018. (in book: Data Analytics: Concepts, Techniques and Applications)

## PROJECTS

---

- ★ Deep learning :: *Malay English Machine Translation System*  
Details can be found here, <https://sbmaruf.github.io/project/mt-system/>
- ★ Deep learning :: *Named Entity Recognition System*  
The details can be found here, <https://github.com/sbmaruf/named-entity-recognition>
- ★ Machine learning system :: *A CBIR System*  
Details of the project can be found here, <http://103.82.172.44:8080/xmlui/handle/123456789/93>
- ★ Deep learning :: *Zero-Resource Cross-lingual Named Entity Recognition*  
Details of the project can be found here, <https://github.com/ntunlp/Zero-Shot-Cross-Lingual-NER>