# Mohammad Saqib Hasan

msaquibhasan@gmail.com mdshasan@cs.stonybrook.edu (631)-681-5567

### **Education**

## Stony Brook, NY

## **Stony Brook University**

**Aug 2021 – Now** 

• Ph.D. in Computer Science. **CGPA**: 3.92. **Research advisor**: Niranjan Balasubramanian. **Research area**: Synthetic Distillation, Preference Optimization, Retrieval Augmented Generation

Dhaka, Bangladesh BUET Jul 2014 – Oct 2018

• B.Sc. in Computer Science and Engineering. CGPA: 3.80.

## **Employment**

## **Applied Scientist Intern**

#### Amazon

May 2025 – Aug 2025

- Worked in the Delivery Experience AI team to develop LLM multi-agent architectures for knowledge augmented database analysis and structured knowledge creation.
- My system helped Amazon engineers save valuable hours in terms of analysis formulation and communication overhead with business management.

Research Engineer BUET Nov 2018 – Dec 2020

• Worked on multiple projects on neural networks ranging from developing to mammal inspired weight initialization to building data optimized (4% of dataset) fake news detectors to neural compression algorithms achieving 5-35x compression. Also won multiple 'ICT Innovation Fund' research grants amounted to \$ 36,000.

# **Technical Experience**

# **Paper**

- Teaching an Old LLM Secure Coding: Localized Preference Optimization with Distilled Preferences ACL 2025
- Handling Open-Vocabulary Constructs in Formalizing Specifications: Retrieval Augmented Parsing with Expert Knowledge. COLM 2024
- Compressed neural architecture utilizing dimensionality reduction and quantization. Applied Intelligence, Springer
- Truth or lie: pre-emptive detection of fake news in different languages through entropy-based active learning and multi-model neural ensemble. ASONAM 2020
- Neuro-scientific analysis of weights in neural networks. IJPRAI, World Scientific

# **Projects**

- Efficient Named Entity Recognition (NER): Devised a framework for NER by combining weak supervision and active learning that achieved full supervision accuracy with only 40% data.
- Analysing Covid Vaccination Rates: Analysed Covid-19 vaccination rates among US counties using public data and social media and created prediction models to measure future rates with low error rates of 1%.

## **Service and Awards**

- Worked on the DARPA SciFy project for Scientific Claim Verification
- Worked on grant NAIRR210140 awarded by NAIRR Pilot Program for developing efficient LLM agents
- Received the prestigious SUNY RF Academic Fellowship for Summer, 2022
- Organizer of the HerWill Datathon 2022.
- Reviewer for journals and ACL Rolling Review (ARR), Neural Processing Letters (Springer Nature) and Artificial Intelligence (Springer Nature)