MOHAMMAD SAQIB HASAN

Email: msaquibhasan@gmail.com | **LinkedIn:** saqib-hasan-a4b326121 | **Github:** SaqibHasan057 | saqibhasan057.github.io/website/

SUMMARY

Ph.D. student in computer science working on creating robust machine learning algorithms for low-resource settings in the semantic parsing domain of natural language processing.

WORK EXPERIENCE

STONY BROOK UNIVERSITY | Research Assistant (RA) — Stony Brook, NY | May 2022 – Present

- Conducting research in Natural Language Processing (NLP) in the Language Understanding and Reasoning (LUNR) lab under Professor Niranjan Balasubramanian.
- Research projects entail auto-formalization of system specifications, hot-started active learning via large language models and semantically similar code scoring.
- Current major work involves developing a transformer architecture model that can incorporate previously unseen domain expert knowledge, provided through means of a lexicon "incontext", during inference for better translation of specification statements into formal language constructs.

STONY BROOK UNIVERSITY | Teaching Assistant (TA) — Stony Brook, NY | Aug 2021 - May 2022

- Acted as lead TA for the course CSE 354 (Natural Language Processing) for the Spring 2022 semester. My duties involved managing other TAs, developing and grading assignments, creating course content and hosting office hours.
- Graded and hosted office hours for the course CSE 310 (Computer Networks) for the Fall 2021 semester.

BUET | Research Engineer — Dhaka, BD | Nov 2018 – Dec 2020

- Worked in the Applied Machine Learning Lab and Cloud Analytics Lab under Professor Muhammad Abdullah Adnan. My tasks involved running separate research projects and writing and presenting proposals for research grants.
- Presented and won the "ICT Innovation Fund" research grant for the project titled **"Detection of Fake News Using Deep Learning for a Cleaner and Safe Internet"** funded by ICT Division, Government of the People's Republic of Bangladesh.
- Worked on two other research grants titled **"Biologically Motivated Learning Rule for Artificial Feedforward Neural Networks"** and **"Trend Detection Using Real Time Principal Component"**.

EDUCATION

PhD in Computer Science — Stony Brook, NY | Aug 2021-Present
Stony Brook University, CGPA: 3.89

B.Sc. in Computer Science and Engineeering — Dhaka, BD | Jul 2014-Oct 2018
Bangladesh University of Engineering and Technology (BUET), CGPA: 3.80

PUBLICATIONS

1. TRUTH OR LIE: PRE-EMPTIVE DETECTION OF FAKE NEWS IN DIFFERENT LANGUAGES THROUGH ENTROPY-BASED ACTIVE LEARNING AND MULTI-MODEL NEURAL ENSEMBLE
   Paper published in **International Conference on Advances in Social Network Analysis and Mining, ASONAM 2020** where I developed and applied uncertainty based active learning to reduce the burden of annotation in supervised methods of automated fake news detection. My methodology proved extremely successful, being able to achieve supervised learning levels of accuracy across benchmark fake news datasets with only **4-28%** of original annotated datapoints.

2. NEURO-SCIENTIFIC ANALYSIS OF WEIGHTS IN NEURAL NETWORKS
   Publication in **International Journal of Pattern Recognition and Artificial Intelligence (IJPRAI)** in which I created biologocially motivated neural network parameter initialization methodologies with the aim of faster training and convergence. By initializating model weights using lognormal values, a statistical distribution very common among the synaptic distribution of the brain, we were able to show a positive correlation with model performance.

3. COMPRESSED NEURAL ARCHITECTURE UTILIZING DIMENSIONALITY REDUCTION AND QUANTIZATION
   Paper in the journal **Applied Intelligence, Springer** for which I developed compression algorithms that combine dimensionality reduction and clustering to create a new architecture with decreased parameterization for neural networks. Our method, when applied on a variety of benchmarks, showed excellent performance with compression as high as 5 to 35 times of the original model with little to no fall in model accuracy.

PROJECTS

1. NLP-ASSISTED FORMAL VERIFICATION OF THE NFS DISTRIBUTED FILE SYSTEM PROTOCOL
   National Science Foundation (NSF) project where I am developing algorithms that utilize large language models (LLMs) and Human-in-the-loop (HIL) methodologies to translate natural language specifications in the Network File System (NFS) Request For Comments (RFCs) into statements in formal modelling languages such as Signal Temporal Logic (STL).

2. WEAKLY SUPERVISED ACTIVE LEARNING SCENARIO FOR NAMED ENTITY RECOGNITION IN DIFFICULT ANNOTATION SETTINGS
   Project where I devise a new training scenario for the Named Entity Recognition task by combining weakly supervised pre-training followed by transfer and active learning of this pre-trained model to train on lower amounts of fully labelled data in a supervised setting, thereby decreasing annotation cost.

3. UNDERSTANDING THE VARIABLES AFFECTING COVID-19 VACCINATION RATES ACROSS COUNTIES IN THE UNITED STATES
   Project where I analyse the variables affecting the Covid-19 vaccination rates across the United States on a county level, thereby understanding interesting trends in the vaccination rates in the US using publicly available socio-economic, medical and political data connected to vaccination. I also evaluated tweets related to the vaccine which are posted by users across the U.S and developed a variety of feature-based machine learning models to see how well I can predict the vaccination rates.

4. EDUENGINEER
   A mobile application which provides a platform for junior IEEE members to share information about research and programming. Developed as part of IEEEmadC 2016 contest.

5. POSTURE CORRECTOR
   An Arduino and Android based application system that can detect improper sitting of a person and notify them via messages through the attached mobile application. Devised with the aim of reducing backpain problem as a result of sitting in irregular positions for a long time.

6. CRIME WATCH
   A mobile application where users can report and learn about crimes and respective statistics pertaining to their area.

ACHIEVEMENTS & HONORS

- SUNY RF Academic Fellowship for Summer, 2022
- Dean's List Award for 2016 and 2018 at my undergraduate alma mater, Bangladesh University of Engineering and Technology (BUET)
- Champion at Hackathon for Environmental Migrants in Bangladesh, organized by Dr. Ingrid Boas, Assistant Professor at the Environmental Policy Group, Wageningen University
- Top 20 at HULT Prize, 2016, organized at Bangladesh University of Engineering and Technology (BUET)

SKILLS

Languages (skilled): Python, Java, C, C++
Languages (known): SQL, Assembly x86, MATLAB, HTML, CSS, Javascript
Technology: Git, Linux, MySQL, LaTeX, Pytorch, Tensorflow, Django, Pandas, AWS, Google Colab, Ionic
Hardware: Arduino, ATMega