Nazia Tasnim

📍 Boston, MA, USA ✉ Email ♥ website 🎓 Nazia Tasnim ⚙ appledora

EDUCATION

GRADUATE
Computer Science, BostonU
Advisor: Dr. Bryan Plummer
2023-present | Boston, MA.

UNDERGRADUATE
CSE, Shahjalal University
of Science and Technology
GPA: 3.97/4.00
2017-2021 | Sylhet, Bangladesh.

RELEVANT SKILLS

Pytorch, Keras, OpenCV, PySpark, Flask, Java

SELECTED PROJECTS

MWTOKENIZER ⚙ i
MWPARSERFROMHTML ⚙
SHAAKO ⚙ i

SELECTED AWARDS

- **CRA-WP** Grad Cohort Fellowship
- **BU GSAS** Dean's Fellowship
- Second Runner-up in **CVPR AI City Challenge** Track-4

LEADERSHIP EXPERIENCE

- Member of Judge Panel, - **BDOSN - NLP Hackathon, 2023**
- Founding Member and Membership Chair - **SUST ACM Students' Chapter, 2019-2020**
- Organizing Committee -**IEEE International Conference on Bangla Speech and Language Processing (ICBSLP), 2019**

REFERENCE

DR. SAJJADUR RAHMAN
Senior Research Scientist |
Megagon Labs
444 Castro St 900, Mountain View, CA 94041

DR. FARIG SADEQUE
Assistant Professor | BRAC University
66 Mohakhali, Dhaka, Bangladesh.

SELECTED PUBLICATIONS

OOD-SPEECH: A LARGE BENGALI SPEECH RECOGNITION DATASET FOR OUT-OF-DISTRIBUTION BENCHMARKING
FR Rakib & SS Dip, Samiul Alam, Nazia Tasnim et al.
INTERSPEECH, 2023

VISTA:VISION TRANSFORMER ENHANCED BY U-NET & IMAGE COLORFULNESS FRAME-FILTRATION FOR RETAIL CHECKOUT
Nazia Tasnim & MIH Shihab & H Zunair, et al.
Proceedings of IEEE/CVF conference on CVPR, 2022

ON LEVERAGING DATA AUGMENTATION AND ENSEMBLE TO RECOGNIZE COMPLEX NAMED ENTITIES IN BANGLA
Nazia Tasnim & MIH Shihab, et al.
International Workshop on SemEval, 2022

WORK EXPERIENCE

IVC-ML GROUP, BU | Research Assistant
Sep 2023 - Present

- Developing probes to analyze biases in large VLMs.
- Building schemes for efficient continual learning across modalities.

WIKIMEDIA FOUNDATION | Research Developer
Sep 2022 - June 2023 | SF, USA.

- Collaborated with Wikimedia's research team to develop NLP tools for 300+ languages in 100+ wikiprojects
- Created extensive research pipeline to curate resources from wiki projects, develop assets through analysis, build models and off-the-shelf tools to be used across all wiki research

GIGA TECH LTD. | Machine Learnign Engineer
Aug 2021 - June 2022 | Dhaka, Bangladesh.

- Developed NLP submodules for the **National Syntactic Treebank for Bangla**
- Created pipelines for **Part-of-Speech (PoS)**, **Named Entity Recognition (NER)** and **language model training**
- Assisted in establishing annotation guidelines for diverse downstream tasks

RESEARCH AFFILIATION

BENGALI.AI | RESEARCH AFFILIATE
PI: Dr. Farig Yousuf
Coordinate teams and lead research projects focused on the NLP landscape of Bangla and other indigenous languages of Bangladesh. We have collaborated with Google, Kaggle and the government of Bangladesh to build some of the largest resources for these otherwise low-resource languages

BRAC UNIVERSITY | RESEARCH ASSISTANT
PI: Dr. Nur Yanhaona
Built an end-to-end pipeline that extensively mines entity information from news reports and produces a knowledge graph, while establishing geopolitical affiliations.

COLLABORATIVE RESEARCH | RESEARCHER
PI: Dr. Sajjadur Rahman & Dr. Farig Yousuf
Exploring the necessity of developing Knowledge Discovery pipelines and resources for Bangla.