sijantanv@gmail.com

Tanvir Ahmed Sijan

https://sijantanvir.github.io/

/tanvir-sijan

/sijantanvir

Research Interests

My research interests lie at the intersection of NLP and AI safety, focusing on human-AI alignment through human-centered evaluation, alignment methods, and LLM interpretability to develop safer and more trustworthy language models.

Education

Jahangirnagar University — Dhaka, Bangladesh
Master of Science in Computer Science — *May 2025 – Sept. 2026*

- **CGPA**: 3.98/4.00
- **Relevant Coursework**: Machine Learning (A+), Deep Learning (A+), Natural Language Processing (A+), Data Mining (A), Artificial Intelligence & Expert Systems (A+), Computational Intelligence (A+)
- **CS Prerequisites**: Data Structures & Algorithms (A+), Software Engineering & Database Systems (A+), Computer Architecture & Operating Systems (A+), Programming Languages (A+)

Bangladesh University of Engineering and Technology (BUET) — Dhaka, Bangladesh
Bachelor of Science in Civil Engineering, with Honors — *Feb. 2020 – Mar. 2025*

- **CGPA**: 3.77/4.00, **Major**: Transportation Engineering, **Minor**: Environmental Engineering
- **Math and Programming Coursework**: Differential Equations & Statistics (A+), Differential, Integral Calculus and Matrices, Computer Programming Sessional (A+), Engineering Computation Sessional (A+)

Test Scores

GRE: 324 (Quantitative: 168/170, Verbal: 156/170, AWA: 4.0/6.0)

Research Experience

Graduate Researcher — Jun. 2025 – Present
Computational Intelligence, Data and Language Research Lab — *Jahangirnagar University*
Supervisor: Dr. Md. Musfique Anwar

- **Mechanistic Interpretability:** Co-conceptualized a study of arithmetic heuristic neurons across three Llama models (1B/3B/8B), testing (1) whether arithmetic neurons are form-invariant across symbolic, python code, and natural language word-problem formulations and (2) whether cross-format failures reflect distinct circuits or differential activation of a shared circuit; conducted experiments on the 1B-parameter model and co-authored the manuscript. Identified a compact, causally sufficient shared circuit whose activation transfer recovered over 97% of cross-format failures for addition and subtraction, with lower recovery for multiplication and division (BlackboxNLP at EMNLP 2026).
- **Event Extraction:** Led a comprehensive study of event detection from Bangla news, developing a generalized news event ontology, coordinating annotation of 9,979 sentences across 40 subtypes and three noise conditions, and training and evaluating encoder-based and decoder-only models; found decoder-only LLMs markedly more robust to noise, while instruction tuning with annotation guideline improved performance on noisy text but did not consistently improve robustness (W-NUT at EMNLP 2026)
- **Cultural Alignment:** Conceptualized and led *BanglaSocialBench* end-to-end, integrating LLM APIs, building evaluation pipelines, coordinating annotation, and conducting statistical analyses for a 1,719-instance benchmark evaluated across 12 LLMs; uncovered systematic over-politeness, failure to recognize multiple socially acceptable address-pronoun forms, and cross-religious kinship-term conflation (ACL SRW 2026, Best Paper Award).

Research Intern — May 2025 – Oct. 2025
Maks Inc. — *Dhaka, Bangladesh (HQ: Frisco, TX, USA)*

- Conducted literature reviews and collaborated with data scientists and transportation engineers to analyze large-scale BRTA traffic datasets, develop emission models, perform geospatial analyses, and conduct in-person stakeholder interviews across major administrative divisions of Bangladesh; prepared technical reports for projects funded by the **World Bank** and **Asian Development Bank**.

Undergraduate Researcher Nov. 2023 – Feb. 2025
Advisor: Dr. Md. Shamsul Hoque; Research Mentor: Dr. Md. Asif Raihan *BUET*

- Built an end-to-end computer vision and machine learning framework integrating street-view imagery, open geospatial data, and crash records to determine the impact of visual and built-environment characteristics on pedestrian crash severity, applying SHAP-based interpretability techniques.

Peer-Reviewed Publications

** indicates equal contribution*

[1] Sharath Naganna*, Tanvir Ahmed Sijan*, Uddipta Kalita. **"Are Arithmetic Heuristic Neurons Form-Invariant? A Mechanistic Analysis of Symbols, Text, and Code in LLMs."** *Proceedings of the 9th BlackboxNLP Workshop: Analyzing and Interpreting Neural Networks for NLP, co-located with EMNLP 2026*, October, 2026. [ACL Anthology] [Code]

[2] Tanvir Ahmed Sijan, S. M Golam Rifat, Nayeemul Islam, Md. Musfique Anwar. **"Beyond Clean Text: Evaluating Encoder and Decoder Robustness for Bangla Event Detection in Noisy Text."** *Proceedings of the 11th Workshop on Natural User-generated Text, co-located with EMNLP 2026*, October, 2026. [ACL Anthology] [Code]

[3] Tanvir Ahmed Sijan, S. M Golam Rifat, Pankaj Chowdhury Partha, Md. Tanjeed Islam, Md. Musfique Anwar. **"BanglaSocialBench: A Benchmark for Evaluating Sociopragmatic and Cultural Alignment of LLMs in Bangladeshi Social Interaction."** *ACL Student Research Workshop*, 2026. **[Best Paper Award]** [ACL Anthology] [Code] [Video]

[4] Tanvir Ahmed Sijan, Md Asif Raihan, Sk. Md. Mashrur, Md. Shamsul Hoque. **"Assessing the Visual and Built Environment Risk on Pedestrian-Vehicle Crash Severity in Absence of Detailed Roadway Inventory: A Deep Learning Based Framework Harnessing Public Geospatial Data."** *Accepted at 106th Transportation Research Board (TRB) Annual Meeting*, Washington, D.C., 2027. [PDF] [Code]

Honors & Awards

Best Paper Award Jul. 2026
ACL Student Research Workshop 2026

Degree Awarded with Honors Mar. 2025
Bangladesh University of Engineering and Technology (BUET)

Undergraduate Student Research Grant Jan. 2024
Research and Innovation Centre for Science and Engineering, BUET

- Awarded competitive student research grant for developing a deep learning based framework, incorporating street-view imagery and open geospatial data with SHAP-based interpretability, to model pedestrian-vehicle crash severity.

Dean's List Scholarship Jul. 2019
Bangladesh University of Engineering and Technology (BUET)

Higher Secondary Certificate (HSC) Scholarship 2019
Board of Intermediate and Secondary Education, Rajshahi

- Awarded Talentpool Scholarship; ranked 10th (∼top 0.01%) among all candidates in the Rajshahi Board.

Service and Leadership

Workshop Reviewer Jul. 2026
11th Workshop on Natural User-generated Text (W-NUT) collocated with EMNLP 2026

Deputy Vice President Dec. 2024 – Mar. 2025
ITE BUET Student Chapter

- Collaborated with executive members to organize workshops and seminars on various topics of transportation research.

Vice President Mar. 2024 – Mar. 2025
ACI BUET Student Chapter

- Organized poster competitions, seminars, and field visits in coordination with chapter members.

Skills

Programming: Python, C, C++, SQL, MATLAB
Libraries: PyTorch, NumPy, Pandas, Transformers, Matplotlib, Scikit-Learn
Other: Linux, Git, Bash, LaTeX
Languages: Bangla (Native), English, Hindi (Basic)