# Tanveer Hossain Munim

Dhaka, Bangladesh · +880 1768-244283 · tanveer.munim@outlook.com · tanveer-munim.site
github.com/munim110 · linkedin.com/in/tanveer-munim · 0009-0000-2295-9325

## Research Interests

**Reliable machine learning:** uncertainty quantification, calibration, and model evaluation.

## Education

**Bangladesh University of Engineering and Technology (BUET)** April 2018 – September 2023
*B.Sc. in Computer Science and Engineering*
**Thesis:** ShasthoSheba, an mHealth service for orphanages (supervisor: Prof. A. B. M. Alim Al Islam)
CGPA 3.62/4.00 · GRE 329 (Quantitative 170, Verbal 159) · IELTS 8.0

## Publications & Manuscripts

### Submitted and under review

[1] **M. T. H. Munim**, B. A. Saiem, A.-A. Sany, and T. Hashem. "A Free Knob: Decoupling Calibration and Predictive Skill in Threshold-Based Evaluation." *Under review at ICLR 2027.* arXiv:2609.33457

[2] **M. T. H. Munim**, B. A. Saiem, and M. A. Adnan. "Calibrate Once, Grow the Vocabulary: Candidate-Independent Scores for Conformal Risk Control in Open-Vocabulary Segmentation." *Submitted to CVPR 2027.*

[3] **M. T. H. Munim**, M. T. T. Anan, F.-Z.-I. Galib, B. A. Saiem, and A. B. M. A. A. Islam. "ShasthoSheba: Leveraging an mHealth Solution for Providing Healthcare Service to Orphanages." *Under review at ACM Transactions on Computer-Human Interaction (TOCHI).*

### Peer-reviewed publication

[4] I. J. Eliza, M. A. Urmi, M. T. T. Anan, **M. T. H. Munim**, F.-Z.-I. Galib, and A. B. M. A. A. Islam. "eDakterBari: A Human-Centered Solution Enabling Online Medical Consultation and Information Dissemination for Resource-Constrained Communities in Bangladesh." *Heliyon*, 10(1), e23100, 2024.

### Work in progress

[5] **M. T. H. Munim**. "Anchored Risk Profiles: Exact Robustness Certificates for Fixed-Threshold Forecast Rankings under Acceptable Corrections." *Manuscript in preparation.*

[6] **M. T. H. Munim**, R. Ahmed, M. A. Karim, K. Lee, and J. Wu. "Confidence-Gated Extraction of Handwritten Bengali Land Records with a Fine-Tuned Vision-Language Model." *Industry-track manuscript in preparation.*

[7] **M. T. H. Munim**, F. Z. Ikram, and A. M. Hassan. "Click Recentring: Regional Count Intervals from a Few Hand-Counted Tiles." *Manuscript in preparation.*

## Research Experience

**Calibration Confounds in Rare-Event Evaluation** August 2025 – Present
*Supervised by Prof. Tanzima Hashem, BUET*

- Showed that max-pooled CSI conflates calibration with spatial skill. A no-retraining monotone recalibration shrinks CasCast's published CSI gain from 0.160 to 0.034; across 450 SEVIR contrasts the frequency-bias gap tracks the shift ($r = 0.80$) and 51 contrasts reverse sign.
- Released the audit as the open-source `freeknob` Python package; the effect reproduces on crowd density and vanishes on well-calibrated segmentation.

**Regional Uncertainty for Object Counting** July 2026 – Present

- Showed with a headroom audit (oracle reference intervals) that no image-derived signal narrows count-scaled conformal intervals across 12 crowd settings and FSC-147; the room for improvement is a per-image bias.
- Proposed click recentring: 4 hand-counted random tiles per image (about 6% of objects) recentre every other region, keep split-conformal coverage, and narrow dense-region intervals by a median 15% under dataset shift (18% on unseen FSC-147 categories).

**Certified Robustness of Forecast Rankings** 2026 – Present

- Exact rational-arithmetic certificates of whether a fixed-threshold ranking survives every acceptable recalibration; on GraphCast/HRES/GFS rainfall, robust leads fall from 109 to 7 as the budget grows from 0.1 to 2 mm.

## Research Experience (continued)

**Vocabulary Shift in Conformal Open-Vocabulary Segmentation** August 2026 – Present
*Supervised by Prof. Muhammad Abdullah Adnan, BUET; submitted to CVPR 2027*

- Showed that conformal risk-control thresholds for CAT-Seg, SED and FC-CLIP fail when ADE20K grows from 150 to 845 classes (up to 15.4% risk at a 10% target), and converted the released weights, without retraining, to reference-anchored scores that keep one calibration valid at every vocabulary size for under one extra label per pixel.

## Professional Experience

**AI GeoLAB Ltd** April 2026 – Present
*Senior Research Engineer*

- **Handwritten land-record AI:** QLoRA-fine-tuned Qwen2.5-VL-3B on 130K labeled Bengali land-record pages (one 16 GB GPU); field scores of **95–97%** on unseen villages versus 79–81% for the best of 11 baselines (Claude Opus 5.5). A token-probability gate passes 53% of index pages to a quick check at 99.3% page accuracy. Deployed for government use.
- **Cadastral map digitization:** ConvNeXt-V2 U-Net, digit-reader ensemble, and roster-constrained assignment recover and correctly number **91.7%** of plots on 22 held-out sheets. Shipped both models in an offline, human-in-the-loop desktop platform (v0.1.0, pilot-ready).

**Regional Integrated Multi-Hazard Early Warning System (RIMES)** July 2024 – July 2026
*Senior Software Engineer*

- Led backend, data-pipeline, and ML architecture for the GCF-funded **Timor-Leste CDIS** national meteorological platform: 1 TB+/day from 5 NWP models and 2 satellites, ∼8K Airflow tasks/day at 97.6% task success.
- Designed OASIS CAP v1.2 alerting used in two countries, the data unification behind Bangladesh FFWC's live 15-day flood forecasts, and a Bengali hybrid-retrieval livestock advisor running a 7B model on a 4 GB GPU.

**Interactive Cares** Jun 2021 – Aug 2022; Jun 2023 – Aug 2024
*Backend Developer; later Tech Lead*

- Led an eight-person engineering team as the learning platform grew from 8,000 to 100,000 users.

## Talks, Fellowships & Software

| | |
|---|---|
| **Invited talks** | "Climate–Agriculture Risk Modeling: Sri Lanka," South Asian Hydromet Forum, Malé, Maldives, 2026.<br>"Machine Learning for Extreme-Event Detection in Agrometeorology," National Center of Meteorology, UAE, 2025. |
| **Fellowship** | National Science and Technology Fellowship, Bangladesh (2026). |
| **Recognition** | NVIDIA Inception Program, MIRA AI (2025) · Accelerating Asia selection, team member (Top 9 of 500+, 2023) · Champion, Bangladesh Mathematical Olympiad; IMO team shortlist (2016). |
| **Software** | wis2downloader (WMO): distributed Celery refactor with spatial filtering before queueing; merged upstream and used in CDIS. · `pycap-validator` (co-author): private operational CAP v1.2 validation software used by two national meteorological agencies. |

## Technical Skills

| | |
|---|---|
| **Reliability & evaluation** | Conformal prediction (split, Mondrian, weighted, risk control), calibration, bootstrap inference, benchmark auditing, pre-registration, PyMC |
| **ML & vision** | PyTorch, Hugging Face, VLM fine-tuning (QLoRA, full), segmentation, OCR, hybrid RAG, llama.cpp, vLLM |
| **Geospatial & climate** | Sentinel-1, Earth Engine, ERA5, IMERG, GloFAS, WeatherBench2, SEVIR, xarray, GDAL, PostGIS |
| **Systems** | Python, Rust, TypeScript, C++, SQL · Airflow, Celery, Docker, Kubernetes |