



IUT 12th ICT Fest Datathon 2026 RULEBOOK



BrainLab presents অলীকবচন: **Bengali LLM Hallucination Detection Challenge**

at the **IUT 12th ICT Fest 2026**, powered by **Institute of Policy Dynamics**, a data science competition designed to strengthen participants' skills in data analysis, machine learning, and artificial intelligence.

Competition Description

অলীকবচন: **Bengali LLM Hallucination Detection Challenge** is a research focused datathon where participants build systems to detect hallucinations in Bengali LLM responses. Given a prompt and a candidate answer (with or without supporting context), teams must determine whether the response is faithful or hallucinated. The competition features a Kaggle leaderboard followed by an in-person final at IUT, aiming to advance reliable Bengali LLMs and multilingual hallucination detection research.

Eligibility Criteria

- Teams can consist of up to 4 members.
- Participant must be a currently enrolled undergraduate student from any recognized university in Bangladesh.
- To verify student status, each participant must submit a valid student ID card during registration.
- International teams are allowed, but at least one team member must be Bangladeshi.
- Cross-university teams are allowed
- Each participant may be a member of only one team for the duration of the competition
- Recent graduates are eligible if they sat for the HSC (or equivalent) examination in 2020 or later and completed their final undergraduate examinations after December 2025. Supporting documents may be requested for verification.



Participation Process

Fill up the registration form while keeping the following points in mind:

1. Team size: 1-4 members from any university.
2. The team must assign a team leader with whom all sorts of communication will be done via email.
3. This will be a private kaggle competition.
4. Ensure that the email address provided has a valid and **verified** kaggle account associated with it for each team member.
5. The registration fee for each team is 600 tk.
6. The Kaggle link will be emailed to the team lead after registration.

Registration Timeline

- Registration Starts: 19th June
- Registration Ends: 15 July
- Any participant can register in between this time frame

Event Timeline

- Preliminary/Online round: 1 July - 20 July
- Preliminary Round Paper and Notebook Submission Deadline - 21 July
- Final Round Participants Announcement - 23 July
- Final/Onsite round: 25 July

Overall Timeline

Event	Date
Registration Starts	19th June , 2026
Datathon starts	1st July, 2026
Registration Closes	15 July, 2026
Online Round Closes	20th July, 2026
Paper and Notebook Submission	21st July, 2026
Final Round Participants Announcement	23rd July, 2026
Final Presentation and Prize Giving	25th July, 2026

Rules and Guidelines

Team Composition

- Each participant may be a member of only one team for the duration of the competition. Multiple submissions from the same participant across different accounts or teams will result in disqualification of all associated submissions.
- Each participant may enter the competition with only one kaggle verified account.
- Team mergers are permitted up to 5 days before the Phase 1 deadline and must be finalized through Kaggle's team-merger interface. No mergers are accepted after that.
- At least one team member must be reachable at a working email address throughout the competition. Communication about advancement to Phase 2 will be sent to that address; teams unreachable within 24 hours of a Phase 2 invitation forfeit their slot.



Competition Structure

This is a two-phase competition. Read both phases before you start.

- **Phase 1 — Kaggle Leaderboard**

Teams submit prediction CSVs directly to Kaggle. Public and private leaderboards are computed on separate splits of the test set given. Phase 1 closes on the date announced in the Timeline; the private leaderboard is revealed at that time.

- **Phase 2 — Solution Package Review and Onsite Final at IUT**

After Phase 1 closes, the top 30 teams on the private leaderboard are invited to submit a complete solution package. (Specified in the submission requirements section).

Organizers run each package on a held-out fold of the dataset that is not part of the public or private Phase 1 test sets. From these results, the top 15 teams are invited to the in-person final at Islamic University of Technology (IUT), where final rankings are determined by a weighted combination of both phases scores. (Clarified in the scoring system section)

The final competition winners are decided by this combined score, not by the Kaggle private leaderboard alone. A team that leads the Kaggle private leaderboard but does not submit a working solution package by the Phase 2 deadline forfeits its place

Models, Training and Compute Rules

- Publicly available **open-weight models** are allowed. These models must be loadable from Kaggle datasets, the Kaggle Models registry, or Hugging Face mirrors that participants attach as datasets to their notebook during notebook submission.
- The use of **any external API** is strictly prohibited. This includes, but is not limited to, OpenAI, Claude, or any hosted or paid services.
- Fine-tuning of pretrained models is permitted but not on the test dataset.
- All inference must be performed **locally within Kaggle notebooks**.



- Total inference runtime must be under 9 hours on a single P100 or T4*2 GPU. Notebooks exceeding this limit are disqualified.
- Total on-disk model size (all weights combined) must be under 50 GB.
- For Phase 1, teams may generate predictions using any resources they have access to, provided the same predictions can be reproduced by a code-competition-compliant Phase 2 package. If your Phase 1 submissions rely on paid APIs and cannot be replicated offline, you will not be able to advance to Phase 2.

External Data Policy

The use of external data is allowed under the following conditions:

- The data must be publicly available, curated from public sources, or created during the competition runtime.
- All external data must be clearly declared and made public with proper citations.
- External data must **not** include or be derived from any competition test set.

Submission Requirements

Phase 1 Submissions

- Format: a CSV file with exactly two columns, id and label, matching the format of sample_submission.csv.
- label must be 0 (hallucinated) or 1 (faithful) for every row. Missing rows, missing labels, non-integer values, or extra columns cause the submission to be rejected without a score.
- Submissions are limited to 4 per day per team.
- A team may select up to 2 submissions as their final scored submissions for the private leaderboard by the Phase 1 deadline. If no selection is made, the two most recent valid submissions are used.



Phase 2 Solution Package (top 30 teams only)

Teams invited to Phase 2 must submit the followings, by the Phase 2 deadline:

- A **training notebook** (Preferred to be Kaggle-runnable but not mandatory) - If applicable. Submitting this training notebook separately is not mandatory but optional.
- A **runnable inference Kaggle** notebook that reproduces the team's Phase 1 predictions from the raw test set, end to end. The notebook must run within Kaggle's standard code-competition kernel limits: GPU (T4*2 or P100) allocation, and total runtime under 9 hours. Organizers will execute this notebook on the held-out fold; notebooks that do not run to completion in this environment are disqualified from Phase 2.
- A **4-page paper report** (excluding references) describing the approach, the methodology, and any analysis of results. Use ACL, EMNLP, or a similar standard NLP formatting template. The report must be a single PDF.
- Model checkpoints or weights (Must be loadable from kaggle datasets, the kaggle models registry or Hugging Face mirrors that participants attach as datasets to their notebooks)
- Clear documentation (README within the notebook or markdown cells)
- Teams are strongly encouraged to upload their paper to [arXiv](https://arxiv.org/) or github and submit the corresponding link.
- A presentation slide which the team will use to present their approach on the onsite day

- **Re-run Policy:**

All submissions will be rerun on both the public and private datasets after Phase 1 concludes.

The inference notebook must execute within Kaggle runtime limits of under 9 hours (CPU or GPU, as permitted).

- Your inference notebook will be re-run on a separate held-out test set. While the input columns and schema will remain unchanged, do not hardcode assumptions such as the number of rows, row order, sample IDs, or any values specific to the provided

`test.csv`.



Submission Limit:

Each team may submit up to 4 submissions per day. Before submitting, ensure that you have joined your designated Kaggle team.

Fair-Play Conduct

- Individual submissions made by team members to bypass the team submission limit may result in disqualification.
- No leaderboard probing beyond the intended purpose. Using submissions to reverse-engineer test-set labels (e.g., systematic label flipping across submissions to infer ground truth) is prohibited and results in disqualification.
- No collusion between teams. Sharing code, predictions, or models privately between teams before the Phase 1 deadline is prohibited. Publicly sharing code, resources, or discussion via the Kaggle Discussion tab is encouraged.
- No misrepresentation of authorship. Every team member listed on the final submission must have made a substantive contribution to the work. Adding names for the sake of prize distribution is prohibited.
- Dataset issues must be reported publicly. If you believe a specific test item is mislabeled, has a contested ground truth, or contains a curation error, report it on the Discussion tab. Reporting privately to try to gain an advantage is prohibited; reporting publicly benefits everyone and is welcomed.
- Organizers reserve the right to disqualify submissions that violate the spirit of these rules even where no specific clause is broken. This clause exists to catch things we could not anticipate
- No manual labeling of the test set. Any attempt to manually inspect, annotate, or otherwise assign labels to the test set for the purpose of improving submissions is strictly prohibited and will result in immediate disqualification.
- Participants must refrain from any form of plagiarism or unethical practices. All work must be original, and proper credit must be given to external sources if used.



Scoring, Ties, and Disputes

- **Primary metric:** Macro F1 on the HALLUCINATED class (label = 0). This is calculated over the private leaderboard rows only for Phase 1 final scoring, and over the held-out fold for the Phase 2 organizer run.
- **Tie-breaker (Phase 1):** when two teams tie on macro F1, the team with the higher F1 on the C1 cultural-distance subset wins the tie. If still tied, the team whose earliest qualifying submission came first breaks the tie.
- Score disputes must be raised via the Discussion tab or by email to the organizing team within 72 hours of the private leaderboard reveal. Disputes raised later are not considered.
- Organizers will publish a post-competition analysis notebook showing per-band and per-domain performance across all top-30 teams. This is analytical, not adjudicatory — final rankings are set by the weighted scheme in Scoring section

NOTE: The authority has the prerogative to make conclusive decisions at any point during the entire competition duration.

NOTE: Use of common sense is highly recommended. Idiotic and deliberate mistakes like idea sharing between teams or trying to submit anything without joining teams won't be tolerated .

Links and Contacts

Registration form	http://iuticfest26.tech/datathon
Contact for queries	Abdullah Al Jubaer Event head, Datathon 01736-587392
Competition Link	https://www.kaggle.com/t/5c9503557c4c404f846028899ea02ce7
Discord Link	https://discord.gg/agVn3E3Nr
Starter Notebook Link	https://www.kaggle.com/code/sushmit0109/starter-notebook-datathon

Prize Pool Breakdown

Award Category	Prize Money
Champion	50,000
1st Runner Up	35,000
2nd Runner Up	20,000
Best Novelty	10,000
Best Paper	5,000



All participating teams will receive **Participation Certificates**, while members of the **Top 15 finalist teams** will receive **Special On-site Finalist Certificates**.

Scoring System

- Phase 1 Online Round (20% of total ; Derived from private leaderboard ranking)
- Phase 2 (80% of total)
 - Held-out Test Set Score (50% of total)
 - In-person Presentation (10% of total)
 - Documentation / Report (10% of total)
 - Approach and Novelty (10% of total)
- Overall scoring scheme
 - Total = Phase 1 Score + Phase 2 Score
 - Phase 1 Score = 0.20* Phase 1 Private Leaderboard Score Only
 - Phase 2 Score = 0.50*(Held-out Test Set Score) + 0.10*(Presentation Score) + 0.10*(Written Paper Report) + 0.10*(Novelty and Approach)

$$0.20+0.50+0.10+0.10+0.10=1.00$$

Judging Criteria:

- Presentation:
 - Clear explanation of assumptions, models, and decisions
 - Informatics and Visualization of the experiments
 - Novelty of the approach
 - Team's' ability to articulate their contributions
- Q&A:
 - Problem Understanding & Motivation
 - Data & Preprocessing



- Modeling & Methodology
- Evaluation & Metrics
- Novelty & Insights
- Report and Documentation:
 - Methodology and Explanation
 - Data visualization and informatics (if applicable)
 - References for the methodology
 - References for new datasets used (If applicable)
- The judges will evaluate each team's written report and presentation. Based on these evaluations, separate scores will be assigned for the quality of the paper and the novelty of the proposed approach. The teams achieving the highest scores in these categories will receive the **Best Paper** and **Best Novelty** awards, respectively.

Changes to These Rules

Organizers may amend these rules to correct ambiguity, address unanticipated situations, or improve fairness. Any material change will be announced on the Discussion tab with at least 1 day's notice before it takes effect, unless the change is a bug fix or clarification that no reasonable team could argue disadvantages them.

For any sort of questions and confusions please ask in the discord server.