Evaluation Policy - LogicLooM 3.0

Contest Rulebook and FAQs, Brochure & Portal Guide - [Click Here](#)
Discussion group - https://chat.whatsapp.com/G98blgQGgaTEm3vIrj8Cce

All participants must compete as part of teams. Technically, individual participations are considered solo "teams" and will be assigned team IDs as well. **Each team must consist of at least 1 and a maximum of 3 participants.** Teams are encouraged to include participants from any combination of the three levels (Foundation, Diploma or Degree).

Note: All grading policies or point allocations will apply to the team as a whole rather than being assigned to individual members within the team, i.e. **we will be computing team points and not participant wise points.****

There will be a **BONUS Quiz activity**. **Attempt Here**

This can bump up to 30 points of your Round-1 score - capped to 100.

Bonus Quiz score = (0.3). Quiz

Model Points (P) [out of 100]

Each team will submit ONE run submission on the un-labelled test data-1 (best of all predictions, as per the labelled validation set scores) and the working code in the 1st level.

The submission will be processed and **ROUGE-1, 2 and L** scores will be auto-evaluated **on 20% of the test set data**. These metrics are in the range [0,1] and will be scaled to 100 %.

Model Points (P) for a Team on test set-1 = 40% Rouge-L + 40% Rouge-1 + 20% Rouge-2

Additional Bonus Policy

#	Bonus Criteria	Points
Participation Modifier	Teams with 2 or more participants	2
Participant Activeness	Attendance in at least one mentoring session by any one of the team-mates	2

Level-1 Filtering Cut-off (L1)

- Condition-1 (C1): Rouge-1 Score >= **0.5 (i.e. 50%)** on <u>Top 20% test data ONLY</u>

OR

- C2: Total Round1 score {P (max 100) **+ Bonus (out of 34 capped at 30)**} - capped at 100 → converted to %ile (relative performance) and is in top 20% ile range

The N1 teams fulfilling the Level-1 conditions now go for Level-2 filtering.

Level-2 Filtering (L2)

Solutions, Working Codes are reviewed for these N1 teams

Rubrics	Criteria
Solution based	Novelty of the solution (robust techniques, approach, models used, efficiency) will be reviewed.
Working code submission	Well-formatted markdown cells in the notebook justifying the rationale behind every code cell (if code is well structured, logically organized, with appropriate use of functions, clear separation of concerns, consistent coding style, meaningful variable names, sufficient commenting for understandability.)

Level-3 Filtering (L3) - If required, for Tie-breaker

If there're ties between teams, the saved (trained) models are run on the surprise private test set-2 (that features content significantly different in style, genre, or domain of the train,dev,test-1 sets).

Generalization, adaptability and robustness are reviewed & higher rank is assigned to the team which has a better score in this check.

Ranks are accordingly assigned based on the descending order of the team points (capped to 100).

Presentation Round - Finale

Top 15 teams will be required to deploy their models to applications and do the system showcase (5 min presentation) on Day-3.

Finale score will be decided by the Judge.

Rubrics for Finale score:

Novelty (10 points)
9-10: Demonstrates exceptional originality and introduces groundbreaking concepts or methods.
7-8: Shows clear innovation and introduces novel ideas or improvements on existing methods.
5-6: Presents some new ideas, but they are not significantly different from existing solutions.
3-4: Limited novelty; primarily relies on well-established techniques or ideas.
1-2: Minimal to no innovation; mostly a rehash of existing work.
0: No originality; merely repeats known information.

Technical Accuracy (10 points)
9-10: All technical details are accurate and well-explained; no errors.
7-8: Minor technical inaccuracies or omissions; overall mostly correct.
5-6: Some technical inaccuracies or misunderstandings; acceptable but with notable issues.
3-4: Multiple technical errors or misunderstandings; significant inaccuracies.
1-2: Major technical inaccuracies; fundamental misunderstandings of key concepts.
0: Technical details are incorrect or misleading; no understanding of the technical aspects.

Clarity and Organization (10 points)
9-10: Presentation is exceptionally clear, well-organized, and easy to follow, logical flow.
7-8: Presentation is clear and organized with minor lapses in flow or structure.
5-6: Generally clear, but with some organizational issues or unclear sections.
3-4: Presentation is somewhat disorganized; clarity issues affect understanding.
1-2: Disorganized and unclear; significant issues with communication and structure.

0: Presentation is completely disorganized and unclear; fails to convey the message.

Model Performance (10 points)
10: Exceptional metric score, highly robust and stable; consistent performance across various datasets and conditions, performs well on new, unseen data without overfitting.
... so on ...
0: No meaningful F1/ accuracy, fails to perform adequately

Visual and Aesthetic Quality (10 points)
9-10: Outstanding visual design; slides and materials are aesthetically pleasing and enhance the presentation.
7-8: Good visual design; mostly clear and visually appealing with minor issues.
5-6: Acceptable visual design; functional but with some distracting elements or minor issues.
3-4: Poor visual design; slides or materials are cluttered or distracting.
1-2: Very poor visual design; slides or materials are unappealing and hinder understanding.
0: No visual design; slides or materials are completely ineffective.

Engagement and Delivery (20 points)
20: Highly engaging delivery; presenter is confident, dynamic, and keeps the audience captivated.
16-19: Engaging delivery with good confidence and energy; maintains audience interest.
11-15: Adequate delivery; some engagement but with noticeable lapses in energy or confidence.
8-10: Limited engagement; delivery lacks energy or confidence, affecting audience interest.
6-7: Poor delivery; disengaging and lacking in confidence, with significant issues in presentation style.
0-5: Completely ineffective delivery; fails to engage or communicate with the audience.

App deployment (30 points)

Based on App UI-UX, time-complexity (time taken by model to predict a caption when user provides a paragraph as input)

Final Team Score = 80% Team Points + 20% Finale Score (out of 100)

** However, for deciding the achievers, to ensure that only the deserving are awarded, individual points & contribution will be reviewed by the team when required on a case to case basis.

—

Prepared by LogicLooM Core Planning Team
For any queries, feel free to reach out to them.