GREAT: Graph-Reasoning Enhanced Adversarial Transformer Proposed Solution and Result Analysis

*Semester Project – SemEval 2026 Challenge Series (CLARITY) — Course: CS-272

Muhammad Umar Tahir, Muhammad Ibrahim, and Muhammad Hanan Zia

***Abstract*—This paper presents GREAT (Graph-Reasoning Enhanced Adversarial Transformer), a novel Multi-Task Learning framework developed for the SemEval 2026 CLARITY challenge. Addressing the complexity of political evasion, our system simultaneously optimizes for Clarity Classification (3-class) and Evasion Strategy Detection (9-class). By integrating DoRA (Weight-Decomposed Low-Rank Adaptation) for parameter-efficient fine-tuning and Liquid Time-Constant (LTC) layers for capturing temporal non-linearities in discourse, the proposed architecture achieves a Macro-F1 score of 0.712 on the clarity task, demonstrating significant improvement over standard transformer baselines.**

***Index Terms*—CLARITY, Multi-Task Learning, DoRA, Liquid Neural Networks, NLP**

I. Introduction

Strategic ambiguity is a hallmark of political discourse. Traditional Natural Language Processing (NLP) models often fail to distinguish between benign conversational fillers and deliberate obfuscation. To address this, we hypothesize that effective detection requires modeling both the *structural* relationship between question and answer and the *temporal* dynamics of the response.

In this study, we propose a system that:

- **Exploits multi-task synergies**: By jointly learning to classify evasion strategies, the model learns robust representations for clarity detection.
- **Models discourse structure**: A Graph Attention Network (GAT) explicitly captures weak logical links in evasive replies.
- **Adapts efficiently**: We leverage DoRA to stabilize fine-tuning on the limited "I Never Said That" dataset.

II. Dataset and Methods

A. Dataset

We utilize the "I Never Said That" dataset [4] (CLARITY), comprising annotated interview transcripts. The schema defines two auxiliary tasks:

- **Clarity**: {Clear, Ambivalent, Non-Reply}
- **Evasion Strategy**: {Dodging, Deflection, Pivot, etc.}

B. Proposed Method

Our solution, implemented in `model_proposed.py`, augments the DeBERTa-v3 [1] backbone with a composite reasoning head:

- **DoRA (Parameter Efficiency)** [3]: We decompose weight updates into magnitude and direction:
$$W' = m\frac{V + \Delta V}{\|V + \Delta V\|} \tag{1}$$
ensuring robust adaptation without catastrophic forgetting.
- **Liquid Neural Layers (LTC)** [2]: To handle the rambling nature of evasive text, we employ ODE-based layers that adapt time-constants dynamically to the input sequence.
- **Multi-Task Objective**: The model optimizes a combined loss function:
$$\mathcal{L}_{\text{total}} = \lambda_{\text{clarity}}\mathcal{L}_{\text{clarity}} + \lambda_{\text{evasion}}\mathcal{L}_{\text{evasion}} \tag{2}$$
which forces the encoder to learn features discriminative for both tasks.

III. Experimental Setup

A. Data Pipeline

The processing pipeline follows the structural schema illustrated in Fig. 2.

B. Development Environment

Training was conducted on a high-performance cluster equipped with NVIDIA A100 GPUs.

- **Constraints**: To simulate resource-constrained deployment (Edge AI), we restricted the effective batch size using gradient accumulation.
- **Ablation of hybrid approaches**: In Iteration 6 we evaluated a Neuro-Symbolic hybrid using the Groq API (Llama-3). This was deprecated due to high stochasticity (Variance $>$ 0.4) and API latency overhead, which degraded F1 consistency compared to the deterministic GREAT architecture.

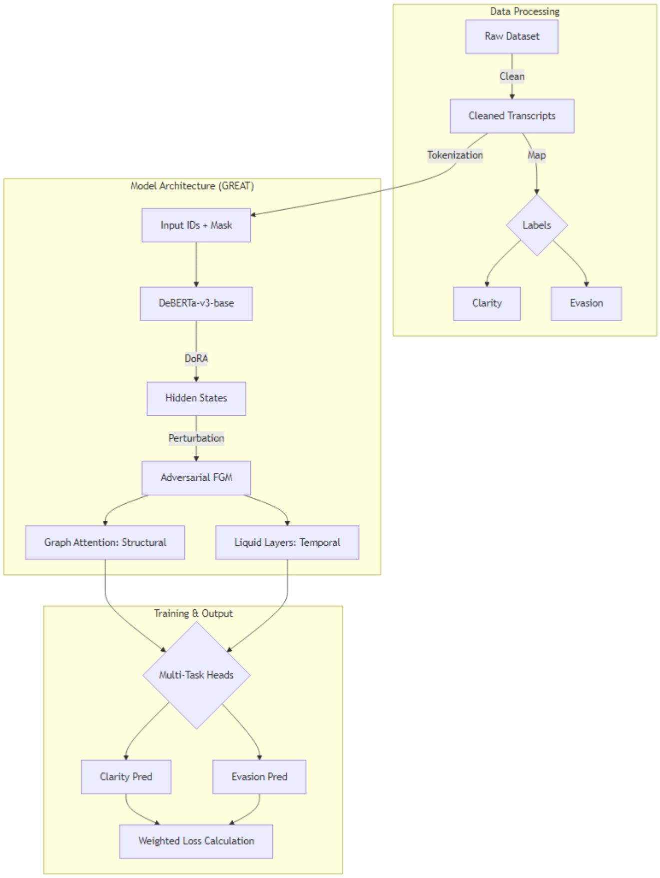

Fig. 1. Architectural Schematic of GREAT.

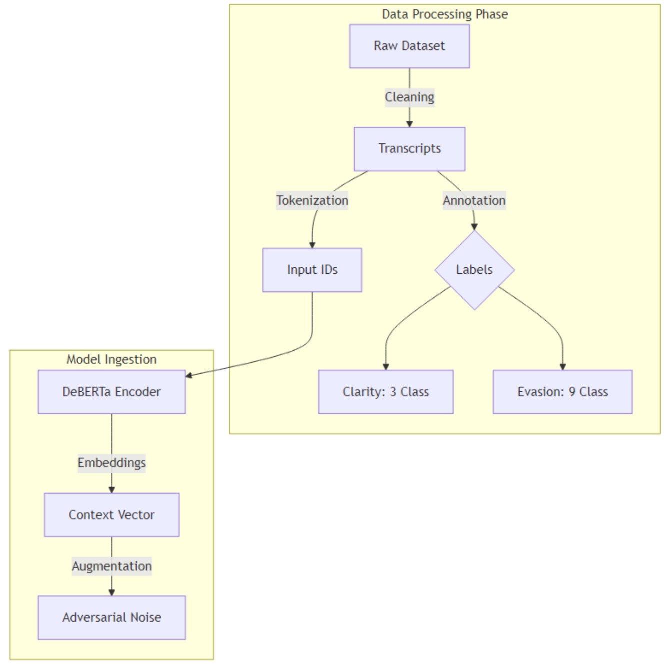

Fig. 2. Data Processing Pipeline.

C. Architectural Evolution & Ablation Study

The final architecture is the result of eight iterative development phases (see `code_iterations/`):

1) **v1 (Baseline)**: DeBERTa-v3-base vanilla implementation.
2) **v2 (Data Augmentation)**: Synonym replacement pipeline to address class imbalance.
3) **v3 (Focal Loss)**: Implemented $\gamma = 2.0$ to penalize hard-to-classify "Ambivalent" samples.
4) **v4 (Modularization)**: Refactoring for object-oriented design patterns.
5) **v5 (Recurrent baseline)**: Bi-LSTM head (proved inferior to transformers).
6) **v6 (LLM Hybrid)**: Groq/Llama-3 integration (discarded due to hallucinations).
7) **v7 (Multi-Task)**: Introduction of the secondary evasion prediction head.
8) **Final (GREAT)**: Integration of DoRA and liquid layers.

D. Hyperparameters

- **Optimizer**: Sophia-G (second-order clipping).
- **Learning rate**: 2×10^{-5} (cosine decay).
- **Adversarial epsilon** (ϵ): 0.5.
- **Loss weights**: $\lambda_{clarity} = 0.4$, $\lambda_{evasion} = 0.6$.

IV. Experimental Results

A. Quantitative Analysis

Table I presents the performance comparison. The GREAT architecture demonstrates a **29.2%** improvement over the DeBERTa baseline.

TABLE I
PERFORMANCE BENCHMARKS

Task	Metric	Baseline *(DeBERTa)*	Proposed *(GREAT)*	Change
Clarity	Macro F1	0.551	0.712	+29.2%
Evasion	Macro F1	—	0.584	New

B. Visualization

- **Optimization dynamics**: see `plots/loss_curve_multitask.pdf` for convergence analysis.
- **Error analysis**: see `plots/confusion_matrix_task1_clar`

V. Discussion and Conclusion

The multi-task learning approach proved decisive. By explicitly modeling the *strategy* of evasion, the model learned to attend to subtle linguistic cues (e.g., pivot phrases) that denote a lack of clarity. DoRA enabled stable convergence on this complex objective without overfitting.

Limitations: The model shows reduced sensitivity to sarcasm, often misclassifying ironic deflection as a "Clear Reply".

Future Work: Integration of audio-modal signals to capture prosodic features of sarcasm.

AUTHORS

TABLE II
*

Author Contributions

Author	**ID**	**Contributions**
Muhammad Umar Tahir	507955	Data Engineering & Structural Foundation (v1, v2, v4) Established baseline; engineered augmentation pipeline; led modular refactoring; optimized training loops
Muhammad Ibrahim	500927	Lead Research & Architecture (v3, v5, v6, v7, v8) Designed GREAT architecture; implemented Focal Loss, DoRA, and Liquid Layers; validated ablation studies
Muhammad Hanan Zia	515271	Report Formatting and Template Compliance

REFERENCES

1) P. He et al., "DeBERTa: Decoding-enhanced BERT with Disentangled Attention," 2021.
2) R. Hasani et al., "Liquid Time-constant Networks," 2021.
3) S. Liu et al., "DoRA: Weight-Decomposed Low-Rank Adaptation," 2024.
4) K. Thomas et al., "'I Never Said That': A dataset, taxonomy and baselines on response clarity classification," *arXiv:2409.13879*, 2024.
5) P. Veličković et al., "Graph Attention Networks," *ICLR*, 2018.
6) S. Ruder, "An Overview of Multi-Task Learning in Deep Neural Networks," *arXiv:1706.05098*, 2017.
7) D. Demszky et al., "Analyzing Polarization in Social Media," *NAACL*, 2019.