adithjrajeev at SemEval-2025 Task 10: Sequential Learning for Role Classification Using Entity-Centric News Summaries

Adith John Rajeev
IIIT Hyderabad
adith.r@research.iiit.ac.in

Radhika Mamidi
IIIT Hyderabad
radhika.mamidi@iiit.ac.in

Abstract

There is a high prevalence of disinformation and manipulative narratives in online news sources today, and verification of its informative integrity is a vital need as online audience is highly susceptible to being affected by such propaganda or disinformation. The task of verifying any online information is, however, a significant challenge. The task **Multilingual Characterization and Extraction of Narratives from Online News**, therefore focuses on developing novel methods of analyzing news ecosystems and detecting manipulation attempts to address this challenge. As a part of this effort, we focus on the subtask of **Entity Framing**, which involves assigning named entities in news articles one of three main roles (*Protagonist*, *Antagonist*, and *Innocent*) with a further fine-grained role distinction. We propose a pipeline that involves summarizing the article with the summary being centered around the entity. The entity and its entity-centric summary is then used as input for a BERT-based classifier to carry out the final role classification. Finally, we experiment with different approaches in the steps of the pipeline and compare the results obtained by them.

1 Introduction

There has been a rapid growth in the field of digital media over the past few years, and this has allowed for much easier dissemination and consumption of information. This development has double-edged outcomes: it allows for more diverse content, and for more real-time engagement with customers, yet it also enables the rapid spread of false information and narratives that are made to influence public opinion or spread propaganda. Consequently, there is a dire need of developing automated approaches that can interpret the framing of entities within a news environment.

The SemEval task on **Multilingual Characterization and Extraction of Narratives from Online News**(Piskorski et al., 2025) has been launched to promote research and develop novel approaches to analyse a news environment and characterize possible attempts at manipulation. Mainly, the first subtask, **Entity Framing** is centered around the classification and identification of the roles relevant entities involved from a news article. A model is required to correctly assign one or more roles to each entity solely based on the context provided by the article, and the locations of entities within it.

We implement this task in a pipeline with 2 primary phases: **1) Entity-Centric Summarization**, and **2) Role classification based on the summary**. Rather than feeding a classifier model the entire news article and the entity directly, we compress the news article into a summary that surrounds the entity in the first phase. This enables the classifier in the second phase, which has a limited token capacity, to focus on a condensed and compact input. The modularity of the pipeline also allows each phase to employ distinct models, and we observe the overall performance to reflect the combined contributions of the models operating at each phase of the pipeline. For the first phase we experiment with some models fine-trained on entity-centric summarization, and also with LLMs as their flexibility allows them to easily adapt to a new task. Most of our analyses, however, is focused on the second phase of the pipeline, where we vary the base models and improve on our training methods to obtain better results. Here, our best results are obtained when we perform sequential training on 3 inter-related tasks: 1) main role classification, 2) fine-grained role classification, and 3) contrastive learning to push the inputs towards a fine-grained class description.

2 Related Work

Previous works in this field have focused on persuasion techniques, framing in the news genre, and

Proceedings of the The 19th International Workshop on Semantic Evaluation (SemEval-2025), pages 241–246
July 31 - August 1, 2025 ©2025 Association for Computational Linguistics

on classifying the roles of entities in memes.

For instance, Ziems and Yang (2021) propose an NLP framework to measure entity framing in the specific case of police violence in the US and demonstrates a difference in how the liberal and conservative news sources frame issues. Additionally, the task of detecting the roles of entities in memes has been addressed by Nandi et al. (2022). This study focuses on classifying entities within memes to the categories of hero, villain, victim or other.

These studies play a role in having an understanding of entity framing across different media formats, and shows the importance of analyzing how entities are portrayed to uncover underlying biases and narratives.

3 Dataset

The dataset consists of news articles covering two major topics—the Ukraine-Russia War and Climate Change—collected from various news aggregation sites between 2022 to mid 2024. The dataset spans 5 languages: Bulgarian, English, Hindi, (European) Portugese, and Russian.

Each article is annotated (Stefanovitch et al., 2025) with the locations of the named entities and their corresponding role. The role assigned to each entity follows a two-level taxonomy:

- A **main role**, chosen from Protagonist, Antagonist, or Innocent.
- One or more **fine-grained roles** that provide a more detailed characterization within the assigned main role.

Additionally, the predefined taxonomy of the fine-grained roles contain a detailed description of all 22 fine-grained roles. While additional annotations exist for other subtasks, these are all the relevant ones to the current subtask.

For our experiments, we use the English subset of the dataset, which contains 3 splits: **train**, **development(dev)** and **test**. The test set labels have not been released yet, so all evaluations were conducted on the dev set. Table 1 provides details on the split, and the distributions of classes within each of them.

Split	Total Entities	Main Role Split (in %)		
		Protagonist	Antagonist	Innocent
Train	686	69.53	18.95	11.52
Dev	91	80.22	9.89	9.89

Table 1: Distribution of main roles in the English dataset. The test set labels have not been released.

4 Methodology

Our main approach to the entity framing task is to implement a two-phase pipeline that uses the article and entity to make an entity-centric summary which is then used to classify the role of the entity. This allows for flexibility in model selection at each phase—different summarization techniques can be explored without drastically altering the classification model's architecture. Similarly, improvements can still be made to the classifier without modifying the summarization process.

4.1 Entity-Centric Summarization Phase

The full length article may contain substantial background information, which need not be entirely relevant to the classification of the role of the given entity. Hence, an entity-centric summarization step may help isolate and condense the content into relevant content, improving the classification accuracy.

The full-length article may contain substantial background information, which need not be entirely relevant to the classification of the role of the given entity. Additionally, entities often appear multiple times throughout the article, sometimes under different mentions or as pronouns. An entity-centric summarization step could identify and consolidate all such mentions, potentially even resolving pronouns referring to the entity if the summarizer is sufficiently capable. This focused condensation of context ensures that only the most relevant portions of the article are retained, which can significantly improve the accuracy of subsequent role classification.

In order to perform the entity centric-summarization, there are 2 main approaches that we implemented to compare results.

The first approach relied on CTRLsum(He et al., 2022), which is a model trained to generate controlled outputs, in this case, it being used to generate entity-centric summaries. The model is fed inputs in the format *Entity* => *Article*, and the model has been fine-tuned to return an output that summarizes the article around the entity.

The second approach was to use an LLM to generate the entity-centric summary. For this, Google's Gemini was used via API calls. We prompted the LLM to return the entity-centric summary given the article and the entity, and this was done in a

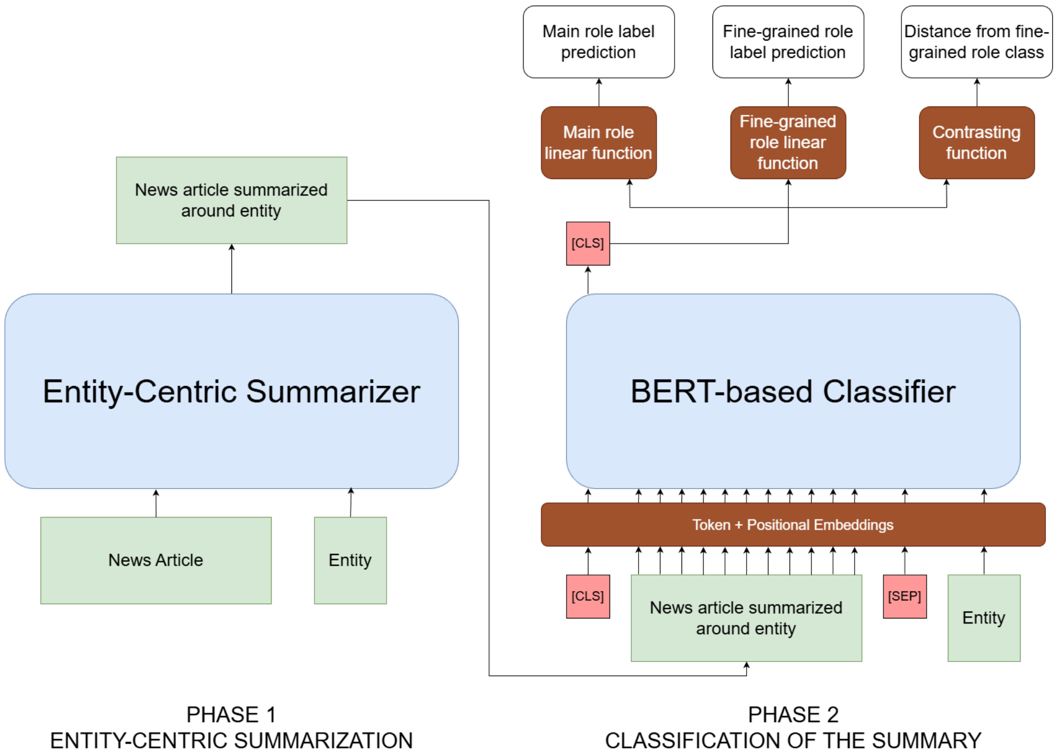

Figure 1: Flowchart of the proposed pipeline.

zero-shot setting. While, the API ensured no compute was happening on our system, we were limited by the free version's number of calls and so it was overall a much slower approach in this stage of the pipeline.

4.2 Role Classification Phase

For the role classification phase of the pipeline, we fine-tuned BERT-based models to predict the main role and fine role with the entity-centric summary and the entity itself as the input. The input was modeled in the following format:

[CLS] Entity-Centric Summary [SEP] Entity

This format leverages BERT's ability to attend to inter-sentence relationships, and allows it to form an understanding of the relation between the summary and the entity, and thereby make a joint representation where both the context and the entity information are fused. This joint representation (via the **[CLS]** token) is then used to classify the entity into its role accurately.

We began by trying a standard BERT architecture with a linear layer on top, that was fine-tuned to predict the fine-role and then mapped to the corresponding main role. The model had a single-cross entropy loss applied uniformly to all role predictions, which provided a first approach for further refinements.

To better capture the structure of the role labels, and to incorporate usage of the main roles, we implemented a dual training strategy. This approach made use of two separate loss functions—one for the main role and one for the fine-grained roles. The training here, happens in a sequential manner as shown below. In each batch in each epoch, the model first predicts the logits for the main loss, and after optimizing for the main loss, the model predicts the logits for the fine loss and then optimizes for it.

Let

- x_b be the input entity-centric summary in the current batch.
- f_θ be the classification model with the parameters θ.
- y_{main} be the ground-truth main role label.
- y_{fine} be the ground-truth fine-grained role labels.

The forward pass consists of two steps:
The model first predicts the logits for the main role:

$$\hat{y}_{main} = f_\theta(x_b) \tag{1}$$

$$L_{main} = CrossEntropyLoss(\hat{y}_{main}, y_{main}) \tag{2}$$

The optimizer updates θ based on the loss L_{main}. After optimizing for the main role, the model predicts logits for the fine-grained roles:

$$\hat{y}_{fine} = f_\theta(x) \tag{3}$$

$$\mathcal{L}_{fine} = BCEWithLogitsLoss(\hat{y}_{fine}, y_{fine}) \tag{4}$$

The optimizer then updates θ based on L_{fine}.

Due to the hierarchical nature of the role classification (from main role to fine-grained role), this sequential training of the two tasks complement one another and potentially allows the model to improve the fine-grained role prediction based on the learnings from the main role prediction.

Following this, we transitioned our backbone model from standard BERT to **DeBERTa v3** for its superior contextual representations and disentangled attention mechanism(He et al., 2020). To address the class imbalance in main role labels, we applied an oversampling strategy by duplicating instances from the underrepresented classes so that all the main roles had an equal number of training examples.

To further improve the model's understanding of the fine-grained roles, we made use of the role descriptions provided in the predefined taxonomy. Each fine-grained role is accompanied by a precise textual definition, which serves as a reference point for classification. Instead of relying solely on the learned label representations from training data, we aimed to explicitly guide the model in aligning the encodings of the input with the that of the fine-grained role descriptions.

This reasoning motivated our final stage of evolution, where we integrated a contrastive loss component into the DeBERTA-based classifier. The contrastive loss was designed to enforce a margin-based separation in the vector space between the positive and negative fine-grained role classes.

Let

- $\mathbf{s} \in \mathbb{R}^d$ be the encoded sentence representation.
- $\mathbf{r}_i \in \mathbb{R}^d$ be the encoded representation of the i^{th} fine-grained role.
- $Sim()$ be a function that returns the cosine similarity between two embeddings.

We first normalize the embeddings:

$$\mathbf{s} = \frac{\mathbf{s}}{\|\mathbf{s}\|_2}, \quad \mathbf{r}_i = \frac{\mathbf{r}_i}{\|\mathbf{r}_i\|_2} \quad \forall i \tag{5}$$

Next we compute the cosine similarity between the sentence embedding and each role embedding:

$$Sim(\mathbf{s}, \mathbf{r}_i) = \mathbf{s} \cdot \mathbf{r}_i \tag{6}$$

For fine-grained roles that are correct(positive pairs), we minimize the negative log-likelihood:

$$\mathcal{L}_{pos} = -\mathbb{E}_{i \in \mathcal{P}} \log \sigma(Sim(\mathbf{s}, \mathbf{r}_i)) \tag{7}$$

For the negative pairs, we enforce a margin m (here 0.5) such that similarity is penalized only when it exceeds m:

$$\mathcal{L}_{neg} = \mathbb{E}_{j \in \mathcal{N}} \max(0, Sim(\mathbf{s}, \mathbf{r}_j) - m) \tag{8}$$

The final contrastive loss is computed as the sum of the positive and negative losses:

$$\mathcal{L}_{contrastive} = \mathcal{L}_{pos} + \mathcal{L}_{neg} \tag{9}$$

We now perform the sequential training for the three tasks: the main role classification, the fine-grained role classification, and the contrastive learning objective for the fine-grained role. This is an extension of approach outlined earlier.

During prediction, an entity is assigned a fine-grained role if it meets both of the following conditions: (1) the fine-grained role classification loss independently predicts it as a positive label and (2) its contrastive similarity score with the role's description falls within the predefined margin. This incorporation and execution of the contrastive loss and sequential training obtained us with the best results in all our experiments.

4.3 Implementation and Hyperparameters

The DeBERTA model 'microsoft/deberta-v3-base' was used for fine-tuning, and different implementations were tried for up to 100 epochs. In our best implementation, we used a batch size of 16. An Adam optimizer was used, with a learning rate of 2×10^{-5} and a weight decay of 0.01.

For the summarizers, we used CTRLsum from the huggingface library, and for the Gemini approach, we used the API function to make calls to

Pipeline		Approach	**Exact Match Ratio**	micro P	micro R	micro F1	Accuracy for main role
Summarizer	Classifier						
CTRLsum	BERT	2-task sequential training	0.13190	0.13190	0.12000	0.12570	0.82420
CTRLsum	DeBERTa	2-task sequential training	0.21980	0.23960	0.23000	0.23470	0.78020
CTRLsum	DeBERTa	3-task sequential training	0.23080	0.25270	0.23000	0.24080	0.82420
Gemini	DeBERTa	2-task sequential training	**0.26370**	0.29900	0.29000	0.29440	**0.83520**
Gemini	DeBERTa	2-task sequential training[1]	**0.25110**	0.32140	0.30570	0.31330	0.86810
CTRLsum	DeBERTa	3-task sequential training[2]	0.23400	0.28090	0.24910	0.26400	0.83400

Table 2: Results from various experiments carried out for the subtask-1.

the Gemini 1.5 Flash model. We had to limit the rate of API calls to under 15 per minute to utilize the free plan of the API.

5 Results and Discussion

As the test set labels have not been released, the analysis has all been done on the dev set and those are the reported scores. Only the final submission has scores for the test set. We have limited our experiments and analysis to the English language.

For each model, the main metrics reported are the Exact Match Ratio, micro P, micro R and micro F1 for the fine-grained roles. An accuracy for the main role prediction is also reported. The official evaluation metric is, however, the EMR for the fine-grained roles. Table 2 shows the various experiments in the order conducted and their performances.

We began with a simple BERT classifier and then moved to a two-task sequential setup using DeBERTa as the backbone model, which brought noticeable performance gains. Further improvements were observed with introducing the third contrastive learning task in addition to main role and fine-grained role classification. We also evaluated the effect of the different entity-centric summarizers: CTRLsum and Gemini.

Our best results were achieved using Gemini as the summarizer in combination with the two-task sequential setup, achieving the highest EMR and micro F1 scores. However, it is worth noting that CTRLsum, despite being a smaller and more efficient summarizer that runs locally (unlike Gemini, which relies on a large LLM, and making API calls), still produced competitive results—especially when combined with the three-task training setup. This makes CTRLsum a practical and resource-efficient alternative, particularly for environments with limited compute resources.

[1]This entry corresponds to the results for the test set, the final submitted entry for the task

[2]This entry corresponds to the results for the test set, using the CTRLsum approach for summarization, experimented after the deadline

6 Conclusion and Future Work

In this work, we explored the effectiveness of a dual phase approach, summarization and classification, in the entity framing task in news articles. For the classifier, we experimented with a sequential training method, incorporating main role classification, fine-grained role classification, and contrastive learning objectives, and this allowed the model to learn hierarchical relationships between main roles and their fine-grained classifications, while the contrastive learning objective helped in creating more discriminative feature representations. While this approach showed promise, achieving an EMR of 0.234, we discovered that the quality of the initial summarization phase had an even more substantial impact on overall performance. Despite the controlled summaries obtained with CTRLsum, our best results (EMR of 0.2637) were achieved using Gemini as the summarizer with a simpler 2-task sequential training setup, possibly due to its higher quality and diversity of summaries.

A key challenge was the data imbalance in both main and fine-grained roles (Table 1), which we addressed through oversampling. However, future work could explore leveraging generative models to create multiple diverse summaries for underrepresented classes for more robust training data.

Our findings highlight the importance of high-quality summarization techniques in framing tasks and suggest promising methods for improving automated narrative analysis in news media.

References

Junxian He, Wojciech Kryscinski, Bryan McCann, Nazneen Rajani, and Caiming Xiong. 2022. CTRLsum: Towards generic controllable text summarization. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 5879–5915, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics.

Pengcheng He, Xiaodong Liu, Jianfeng Gao, and Weizhu Chen. 2020. Deberta: Decoding-enhanced bert with disentangled attention. *CoRR*, abs/2006.03654.

Rabindra Nath Nandi, Firoj Alam, and Preslav Nakov. 2022. Detecting the role of an entity in harmful memes: Techniques and their limitations. In *Proceedings of the Workshop on Combating Online Hostile Posts in Regional Languages during Emergency Situations*, pages 43–54, Dublin, Ireland. Association for Computational Linguistics.

Jakub Piskorski, Tarek Mahmoud, Nikolaos Nikolaidis, Ricardo Campos, Alípio Jorge, Dimitar Dimitrov, Purificação Silvano, Roman Yangarber, Shivam Sharma, Tanmoy Chakraborty, Nuno Ricardo Guimarães, Elisa Sartori, Nicolas Stefanovitch, Zhuohan Xie, Preslav Nakov, and Giovanni Da San Martino. 2025. SemEval-2025 task 10: Multilingual characterization and extraction of narratives from online news. In *Proceedings of the 19th International Workshop on Semantic Evaluation*, SemEval 2025, Vienna, Austria.

Nicolas Stefanovitch, Tarek Mahmoud, Nikolaos Nikolaidis, Jorge Alípio, Ricardo Campos, Dimitar Dimitrov, Purificação Silvano, Shivam Sharma, Roman Yangarber, Nuno Guimarães, Elisa Sartori, Ana Filipa Pacheco, Cecília Ortiz, Cláudia Couto, Glória Reis de Oliveira, Ari Gonçalves, Ivan Koychev, Ivo Moravski, Nicolo Faggiani, Sopho Kharazi, Bonka Kotseva, Ion Androutsopoulos, John Pavlopoulos, Gayatri Oke, Kanupriya Pathak, Dhairya Suman, Sohini Mazumdar, Tanmoy Chakraborty, Zhuohan Xie, Denis Kvachev, Irina Gatsuk, Ksenia Semenova, Matilda Villanen, Aamos Waher, Daria Lyakhnovich, Giovanni Da San Martino, Preslav Nakov, and Jakub Piskorski. 2025. Multilingual Characterization and Extraction of Narratives from Online News: Annotation Guidelines. Technical Report JRC141322, European Commission Joint Research Centre, Ispra (Italy).

Caleb Ziems and Diyi Yang. 2021. To protect and to serve? analyzing entity-centric framing of police violence. In *Findings of the Association for Computational Linguistics: EMNLP 2021*, pages 957–976, Punta Cana, Dominican Republic. Association for Computational Linguistics.