GR-EN-A-DE: A Novel Graph-Based Model for online Extremist Narrative Analysis

Anonymous ACL submission

Abstract

Graph-based modeling has emerged as a prominent approach in hate speech detection and Extremist narratives analysis, offering a flexible framework to encode complex textual, structural, and semantic signals. We propose GR-EN-A-DE, a novel graph-based Extremist Narrative model grounded in established linguistic and sociological theories. On top of this model, we design two learning strategies: a semi-supervised graph encoder–decoder and a contrastive learning framework for extremist feature detection. Our method is evaluated on a large-scale Extremist Narrative classification benchmark that considers various multilingual data (real and synthetic), comparing also our approach to state-of-the-art masked language models.

Figure 1: Example of Extreme Narrative structure carved out from the context of anti-immigrant sentiment spread. *In-Group:* refers to the group in need of defense/protection, and the *Out-Group:* is the target of a negative claim. The *Problematization* is characterized by distrust of institutions and poses a threat to culture and safety, with a *Possible Solution* that involves calling for radicalization.

1 Introduction

Extremist Narratives (EN) are meaning-making systems that structure events, actors, and proposed actions in ways that divide the world into morally superior in-groups and dangerous, inferior out-groups (Postigo-Fuentes et al., 2025). They operate along two main axes: the relationship between in-group and out-group, and the proposed solution to a perceived crisis (Postigo-Fuentes et al., 2025; Ingram, 2016b). Extremism is expressed through the rigidity of the Us/Them dichotomy and the harm potential of the solution, ranging from exclusion and silencing to calls for violence (Mudde and Rovira Kaltwasser, 2017; Määttä, 2023).

In this sense, the annotation schema developed by Baider and Gregoriou (2024) provides an operationalization of EN structure by focusing on extreme speech and its linguistic realization. Such schema identifies not only in-groups and out-groups, but also initiating problems, solutions, emotional appeals, rhetorical devices, and extreme speech parameters such as intolerance, hostility, superiority, polarization, and perceived threat. These features draw on well-established insights into the linguistic construction of hate and extreme speech, such as the emphasis on the in-group's positive qualities and the out-group's negative traits (Van Dijk et al., 1998; Allport et al., 1954; Musolff, 2015), the use of dehumanizing metaphors (Baider and Kopytowska, 2018), and the deployment of conspiracy theories and threat narratives (Udupa et al., 2021; Richardson and Wodak, 2022). To that extent, the schema offers the linguistic and textual granularity necessary to detect how Extremist Narratives are articulated in discourse.

In Figure 1, we depict an example of an Extremist Narrative in the context of an anti-immigrant rhetoric (Veri and Maier, 2025; Smets et al., 2019) spreading in online platforms. In a typical scenario, local and hard-working people (the *In-Group*) feel attacked by immigrants and local government policies. Such a structured relationship is seen to be evolving around a problematization, where culture and safety are in danger, while the proposed solution involves reinforcing, or even radicalizing, anti-imigrant narratives to protect the in-group.

The feature-rich annotation schema proposed by Baider and Gregoriou (2024) enables a nuanced understanding of extremist discourse and creates opportunities and challenges for computational modeling. Typically, narrative classification solutions rely on large pre-trained models (i.e., Transformer-based models) (Vaswani et al., 2023) or traditional ML classifiers. These approaches excel in capturing sequential and contextual dependencies in natural language; however, they are not designed to model and thus leverage heterogeneous, multi-feature datasets that combine textual, structural, and semantic signals.

In this sense, we propose to define and tackle Extremist Narrative analysis by adopting graph representation learning, which offers a principled way to jointly encode multiple feature types and their dependencies through node and edge relations, enabling richer contextualization than purely sequence-based architectures.

Contribution We propose a novel Graph Extremist Narrative Model, called GR-EN-A-DE, designed based on a well-established linguistic and sociological framework of Extremist Narratives across multiple contexts. Upon GR-EN-A-DE, we design two learning strategies: *(a)* a semi-supervised graph encoder-decoder framework, and *(b)* a contrastive learning framework for extremist feature detection. We evaluate our contribution on an extensive EN classification benchmark, considering both real and synthetic datasets and comparing our approach against state-of-the-art masked language models.

2 Related Work

Graph-based modeling has emerged as a prominent approach in hate speech detection and social media analysis, offering a flexible framework to encode complex linguistic, social, and conversational structures. Textual elements such as tokens, posts, or tweets are often represented as nodes, with semantic relations (e.g., PMI (Fano, 1963), TF-IDF (Joachims et al., 1997)) defining edges, and thus, enabling Graph Neural Networks (GNNs) to capture fine-grained contextual dependencies (Wu et al., 2024; Bölücü and Canbay, 2021; Duong et al., 2022). Beyond text, user-centered approaches model social interactions and community structure to incorporate behavioral context, improving robustness and generalization (Mishra et al., 2019, 2018; Del Tredici et al., 2019; Chakraborty et al., 2022). Temporal and structural graph representations of retweet cascades and threaded discussions further support early detection of hate propagation (Beatty, 2020; Hebert et al., 2022, 2023). Advanced architectures, including hypergraphs, heterogeneous graphs, and multimodal models, extend this capacity by capturing higher-order or cross-modal relationships (Mu et al., 2024; Mou et al., 2021; Hebert et al., 2024). Graph-based dialogue models have also been applied to counter-narrative generation by leveraging full conversational context (Baez Santamaria et al., 2024). Furthermore, graph-based studies have examined how platform dynamics shape toxicity and polarization (Chavalarias et al., 2023; Gaumont et al., 2018), while recent work highlights explainability by identifying nodes and relations driving GNN predictions (Wasi, 2023).

Following the research line introduced, we present a novel graph-based model that accounts for the complex relationships in the context of Extremist Narrative analysis. Before formally introducing the model and the Machine Learning framework, we define, in the following section, the narratological aspects and entities at the core of Extremist Narratives.

3 The GR-EN-A-DE Model

3.1 Extremist Narratives Features

Extremist Narrative Identification. The study of Extremist Narratives has recently shifted from theoretical analyses of ideology and discourse to systematic, operational perspectives capable of capturing their structural and contextual complexity, and revealing how Extremist Narratives construct moral hierarchies and collective identities, portraying **in-groups** as superior and **out-groups** as threatening, thereby facilitating polarization and legitimizing hostile or exclusionary actions (Määttä, 2023; Beyer and Schauer, 2021). Such narratives often rely on storytelling, emotional appeals, historical references, and conspiratorial or populist elements to link crises to out-groups and solutions to in-groups, framing radical or violent action as morally justified (Postigo Fuentes et al., 2024; Ingram, 2016a; Reed and Dowling, 2018; Glazzard, 2017; Hidalgo, 2021) .

Extremist Narrative Entities. Based on the introduced conceptual framework and the extensive study of Baider and Gregoriou (2024) and Postigo-

Fuentes et al. (in press), we consider three interconnected levels in the discourse: (I) **Group identity level**, namely *In-group* and *Out-group*, (II) **Thesis level**, the central problem, a.k.a *Initiating problem* attributed to the *Out-group*, and (III) the **Outcome level**, which identifies the target actions and the presence of a radical responses, namely *Hostility*, *Intolerance*, *Threat*, *Polarization*, *Solution* and *Topic*.

In our work, we argue that learning Extremist Narrative patterns must rely on the identification of a relational topology in which meaning emerges through the connection of different contexts. Therefore, rather than treating a corpus under the lens as a collection of independent (annotated) elements, we aim to define a graph structure that explicitly states the narrative logic and the level of connection, e.g., who speaks, about whom, and toward which actions. Intuitively, such a representation can directly capture both the local tensions (e.g., a cluster binding an out-group to a specific grievance) and the broader narrative trajectories that guide audiences from identity claims to action prescriptions.

In the following sections, we formally introduce the proposed model and subsequently describe the novel Extremist Narrative classification framework that leverages it.

3.2 Extremist Narrative Graphical Model

We define a graph model in which in-group and out-group features are encoded as attributes of extremist discourse connections among nodes, which in turn represent textual feature dimensions. Our assumption is that this model captures *potential* causal relationships within the complex rhetoric of extremist discourse, thereby providing a spatial structure for the narrative.

Formally, we define the Extremist Narrative Model (GR-EN-A-DE) as an attributed graph: $\mathcal{G} = (V, E, \phi_V)$ where: (V) is a finite set of nodes, $(E \subseteq V \times V)$ is a set of undirected edges, ϕ_V is a node attribute function.

Each node $v \in V$ is associated with attributes provided by the following functions:

Hostility : $V \rightarrow \{0,1\}$, *Initiating Problem* : $V \rightarrow \{0,1\}$, *Intolerance* : $V \rightarrow \{0,1\}$, *Threat* : $V \rightarrow \{0,1\}$, *Polarization* : $V \rightarrow \{0,1\}$, *Solution* : $V \rightarrow \{0,1\}$, *Topic* : $V \rightarrow \mathcal{T}$, where $\mathcal{T} = \{t_1, t_2, \ldots, t_k\}$ is a finite set of topic categories, *Target Group* : $V \rightarrow \mathcal{TG}$, where $\mathcal{TG} = \{tg_1, tg_2, \ldots, tg'_k\}$ is a finite set of narrative target group, *Annotation Type* : $V \rightarrow \mathcal{AT}$, where $\mathcal{AT} = \{at_1, at_2, \ldots, at''_k\}$ is a finite set of annotation type categories, *In-Group* : $V \rightarrow \mathcal{L}$ and *Out-Group* : $V \rightarrow \mathcal{L}$, where $\mathcal{L} = \{g_1, g_2, \ldots, g_k\} \cup \{\perp\}$ be the finite set of group identities of interest, with $\perp$ denoting the undefined value.

Therefore, the node attribute function is defined as: $\phi_V(v)$ = *Hostility*(v), *Initiating Problem*(v), *Intolerance*(v), *Threat*(v), *Polarization*(v), *Solution*(v), *Topic*(v), *Target Group*(v), *Annotation Type*(v), *In-Group*(v), *Out-Group*(v), where $v \in V$.

4 Extremist Narrative Classification Framework

Here, we present the details of our novel EN classification framework based on GR-EN-A-DE. In Figure 6 we depict the main algorithmical components, showing the overall architecture of our EN classification framework composed by three different blocks : I) Embedding computation and graph initialization, II) Graph Model Learning, and III) Node attribute prediction.

4.1 Graph Initialization

In our solution, we start to construct $\mathcal{G}$ considering that each vertex $v_i \in V$ refers to a text sample embedding represented in a d-dimensional feature space ($\mathbb{R}^d$).

We illustrate in Figure 3 the GR-EN-A-DE construction steps. We initialize the graph using the K-nearest neighbor algorithm (Cover and Hart, 1967) on the node features, generating an initial set of edges ($E_{kNN} \subseteq E$), which connect the K similar nodes to each candidate. We enrich the model, adding extra edges ($E_{att} \subseteq E$) between vertices that encode text concerning the same In-group or Out-group. Hence, for each $(v, v') \in E_{att}$, at least one of the following conditions hold: *In-Group*(v) = *In-Group*(v') and *Out-Group*(v) = *Out-Group*(v').

We leverage the topology of the generated graphical model using two graph learning paradigms, presented in the following subsection.

4.2 Graph Model Learning

Semi-supervised graph learning Semi-supervised graph learning refers to a class of learning paradigms in which models leverage both labeled and unlabeled nodes in a graph to infer node representations and predict target properties (Jiang and Bai, 2024), exploiting the relational structure

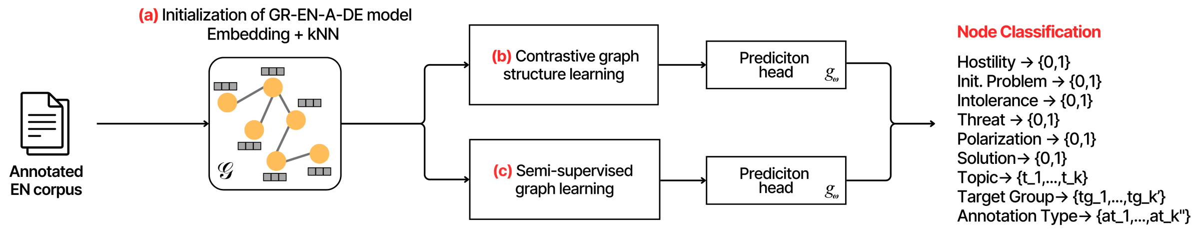

Figure 2: Extremist Narratives Classification Framework. *(a)* Initialization of GR-EN-A-DE model from the annotated corpus. *(b)* Self-supervised contrastive graph structure learning. *(c)* Semi-supervised graph learning *(d)* Masked Language Model baseline.

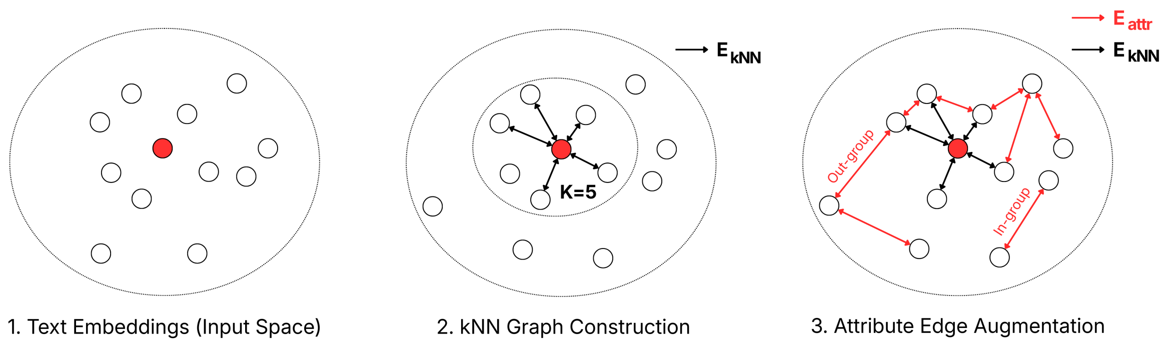

Figure 3: EN Graph construction pipeline. (1) Each text smaple is embedded using a pretrained SBERT model and represents a node. (2) We compute kNN edges for each node based on embedding similarity and construct an initial graph. (3) We add co-membership edges based on shared EN categorical attributes (In-Group, Out-Group, etc.). The final graph captures both semantic and contextual relationships among text samples.

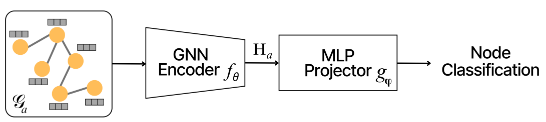

Figure 4: Semi-supervised graph learning

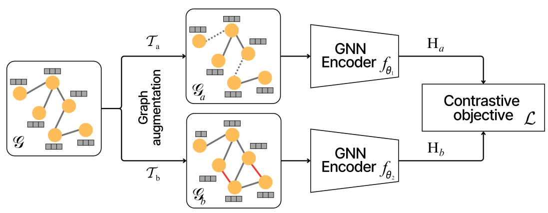

Figure 5: Self-Supervised graph contrastive learning. The input $\mathcal{G}$ undergoes two data augmentations τ_a and τ_b to generate the two graph views $\mathcal{G}_a$ and $\mathcal{G}_b$, both of which are encoded by the same GNN f_θ to produce the representations H_a and H_b, which are then optimized using the contrastive loss $\mathcal{L}_{contrastive}$ in order to maximise the agreement between the two views.

encoded in the graph, such as edges representing interactions, similarities, or relations between node attributes.

In Figure 4 we show the architecture of semi-supervised learning. For a given graph $\mathcal{G} = (V, E)$ with node features $\mathbb{R}^{|V|\times d}$, the encoder is defined as: $H = f_\theta(X, A)$ where A is the adjacency matrix of the graph, $H \in \mathbb{R}^{|V|\times d'}$ is the matrix of node embeddings (with $d' < d$), and f_θ is a graph neural network parameterized by θ that aggregates neighborhood information. The learned embeddings are then passed through a task-specific MLP head for prediction $\hat{Y} = g_\phi(H)$, where g_ϕ is a multi-layer perceptron parameterized by ϕ and $\hat{Y}$ contains the predicted attributes of the nodes.

Contrastive Learning Contrastive Learning has emerged as a powerful framework for graph representation learning, whose objective is to align node embeddings across multiple augmented views of the same graph (You et al., 2020; Liu et al., 2022). Although originally designed for unlabeled graphs, we extend contrastive learning to incorporate available node labels, to guide embedding refinement and improve representation quality.

Our approach builds upon the **StrUcture Bootstrapping** contrastive **LearnIng** fra**MEwork** (**SUBLIME**) proposed by (Liu et al., 2022), which effectively captures structural information through graph augmentations and contrastive objectives.

As illustrated in Figure 5, the input graph $\mathcal{G} = (V, E)$, with node features $X \in \mathbb{R}^{|V|\times d}$, is subjected to two stochastic augmentation functions, τ_a and τ_b, resulting in two distinct graph views $\mathcal{G}_a$ and $\mathcal{G}_b$. These augmentations can involve random edge/node dropping, attribute masking, or subgraph sampling. Each view is then passed through a shared graph neural network (GNN) encoder f_θ, yielding node representations $H_a = f_\theta(X_a, A_a)$ and $H_b = f_\theta(X_b, A_b)$, where A_a and A_b denote the adjacency matrices of the respective views.

To encourage view-invariant representations, we optimize a contrastive loss $\mathcal{L}_{\text{contrastive}}$ that aligns embeddings of the same node across views while pushing apart representations of different nodes.

The contrastive loss is defined as:

$$\mathcal{L}_{\text{contrastive}} = \frac{1}{2n}\sum_{i=1}^{n}\left[\ell(z_{a,i}, z_{b,i}) + \ell(z_{b,i}, z_{a,i})\right],$$

where

$$\ell(z_{a,i}, z_{b,i}) = -\log\frac{\exp\left(\text{sim}(z_{a,i}, z_{b,i})/t\right)}{\sum_{k=1}^{n}\exp\left(\text{sim}(z_{a,i}, z_{b,k})/t\right)}.$$

Here, $\text{sim}(\cdot,\cdot)$ denotes cosine similarity, t is a temperature parameter, and n is the number of nodes.

Node embeddings are enriched by concatenating textual features x_i with additional encoded node attributes c_i (i.e., in-group and out-group), producing a combined representation

$$z_i = f_\theta(x_i, c_i)$$

that captures both structural and contextual information. These enriched embeddings are then used in the contrastive objective, where the similarity $\text{sim}(z_{a,i}, z_{b,i})$ reflects both graph structure and node-level context, strengthening the quality of the learned representations for downstream node classification tasks.

5 Experimental Setup

In this section, we present the experimental framework employed to evaluate the proposed EN classification framework, detailing the datasets and model configurations.

Model training and evaluation were carried out using key libraries such as `transformers`, `datasets`, `PyTorch-geometric`, and `DGL`[1]. Comprehensive information on package versions and the computational environment is available in our repository to facilitate reproducibility: https://anonymous.4open.science/r/ARR_PPR_SUBMISSION-8139/.

[1] https://www.dgl.ai/

5.1 Datasets and Data Preparation

To evaluate our approach, we use three datasets covering multiple languages and containing annotations and metadata related to Extremist Narratives. These include **ToxiGen** (Hartvigsen et al., 2022), **FRENK** (Ljubešić et al., 2019), and a newly collected **Multilingual EN Corpus** curated from social media platforms in French, German, Slovenian, and Cypriot Greek.

The first two datasets are publicly available under permissive licenses, and the Multilingual EN Corpus is intended to be released for research use. Our use of these resources is consistent with their original purpose, primarily classification and moderation of hate and extremism related content.

In the following subsections, we describe each dataset and its preprocessing steps in detail. We moreover add in the Appendix (section A) finer details of the analyzed datasets.

5.2 Models and Training Setup

We evaluate two graph based modeling approaches, namely a self supervised graph learning framework and a contrastive graph learning framework, and compare them with a transformer based text classification baseline using BERT Base uncased (Devlin et al., 2019). All experiments are conducted on an infrastructure equipped with NVIDIA A100 GPUs with 80 GB of memory, 2 TB of system memory, and dual AMD EPYC 7543 Milan processors with 32 cores operating at 2.80 GHz.

5.2.1 Framework Settings

To build the GR-EN-A-DE model (refer to Figure 6), we first embed text instances using a pretrained sentence encoder (Sentence Bert) (Reimers and Gurevych, 2019).

Semi-Supervised Learning In the Semi-Supervised Learning pipeline, we employ the GraphSAGE framework (Hamilton et al., 2018), which adopts a graph encoder to learn node embeddings using the iterative message passing paradigm, allowing information to diffuse across semantically similar nodes and mitigate noise in individual feature vectors. To optimize the hyperparameters, we conduct a grid search over dropout rates (0.2, 0.4, 0.6), hidden dimensions

(64, 128), learning rates (1e-3, 5e-3), activation functions (relu", "elu"), and Jumping Knowledge models (JK) ("lstm", "cat", "max") following (Xu et al., 2018). Each configuration was trained over 600 epochs on a fixed random seed (42). We manually inspect the loss values evolution during the training of all these parameters and run ablation studies to confirm the behavior of each one. We notice that the configuration with the highest dropout (0.6) and a large lr (0.005) achieved the best validation performance overall. Upon close inspection we realize that this happens only after two epochs suggesting that the model quickly learned to separate the classes but may overfit afterwards. This is why for the experiments we employ a more conservative setup (dropout 0.2, lr 0.001, ELU, JK-LSTM) which showed a slower and more stable learning.

Contrastive Learning In the Contrastive Learning pipeline, the views are refined using Feature Graph Prior (FGP) learners Liu et al. (2022). Node embeddings are optimized via contrastive learning with random edge dropout ($p_a = 0.2$) and feature masking ($p_x = 0.35$). Both views are encoded with a shared 2-layer GraphSAGE encoder (Hamilton et al., 2018) (hidden dimension 256, dropout 0.5) and projected through an MLP (representation dimension 256, projection dimension 128). Training runs for 4000 epochs using Adam (learning rate 0.005, weight decay 1×10^{-4}) with a global contrastive loss temperature of $t = 0.08$.

Downstream Node Classification Node embeddings are used for classification with a two-layer GraphSAGE model (hidden dimension 256, dropout 0.5), incorporating Batch Normalization and Jumping Knowledge concatenation to aggregate information across layers. A fixed train/validation/test split (60/20/20) is applied, and models are trained for 600 epochs using Adam (learning rate 0.01, weight decay 5×10^{-4}). Performance is evaluated using macro-averaged F1 of classification performed on the test set . Multiple seeds (0, 21, 42, 84, 123) are used to ensure robustness, and results are reported as mean ± standard deviation.

6 Results

6.1 Sensitivity Analysis

We first conduct a sensitivity analysis to assess the impact of graph construction choices on node classification performance. In particular, we vary the number of neighbors $k \in \{5, 10, 15, 25\}$ used in the k-nearest neighbor graph and the maximum number of attribute edges per node in E_{attr} denoted as Max EN $\in \{0, 10, 50, 100, 150, 200\}$. When Max EN is equal to 0, the considered graph is obtained exclusively by applyinh the $K-NN$ algorithm.

6.1.1 Discussion

In Figures 6 and 7 we present the results of the sensitive analysis for all the datasets, varying the Max EN and K respectively.

Detailed results for each dataset are provided in the appendix (Tables 2, 3, 4, 5, 6, 7, 8, and 9, 10, 11, 12, 13).

Toxigen Incorporating a small number of auxiliary features substantially improves performance on Toxigen across both learning paradigms (see Tables 2, and 3). For **Semi-Supervised** graph learning, the results indicate that performance is mainly influenced by the maximum number of auxiliary features rather than the choice of k. For example, at $k = 5$, increasing Max from 0 to 10 improves the score from 68 to 72. Beyond this point, gains diminish and performance stabilizes or slightly decreases, while variations in k lead to only marginal changes. When using **Contrastive Learning**, a similar trend is observed: Max EN = 10 achieves the highest F1 scores (65-66), outperforming Max EN = 0 (63-64) by approximately 1.5-2 points. Further increases in Max EN provide only marginal or inconsistent improvements, and changes in k have limited impact, indicating that a small amount of auxiliary information captures most of the benefit.

FRENK Incorporating auxiliary features substantially improves performance on the FRENK dataset across both learning paradigms (see Tables 4, 5). For **Semi-Supervised** graph learning, performance is strongly influenced by the number of node features (Max EN), while variations in the number of neighbors k have a smaller effect. For example, in the LGBT subset at $k = 10$, increasing Max EN from 0 to 10 decreases the F1 score from 99.17 to 86.33, but further increases in Max EN (50–200) recover and stabilize performance around 99.0–99.1. For the Migrants subset, scores are already high at Max EN = 0 (98.95) and remain stable across higher Max EN values. **Contrastive Learning** shows a stronger positive effect of auxiliary features. In FRENK-LGBT, F1 scores increase from 83–90 at Max EN = 0 to 97–99 for Max EN ≥ 10,

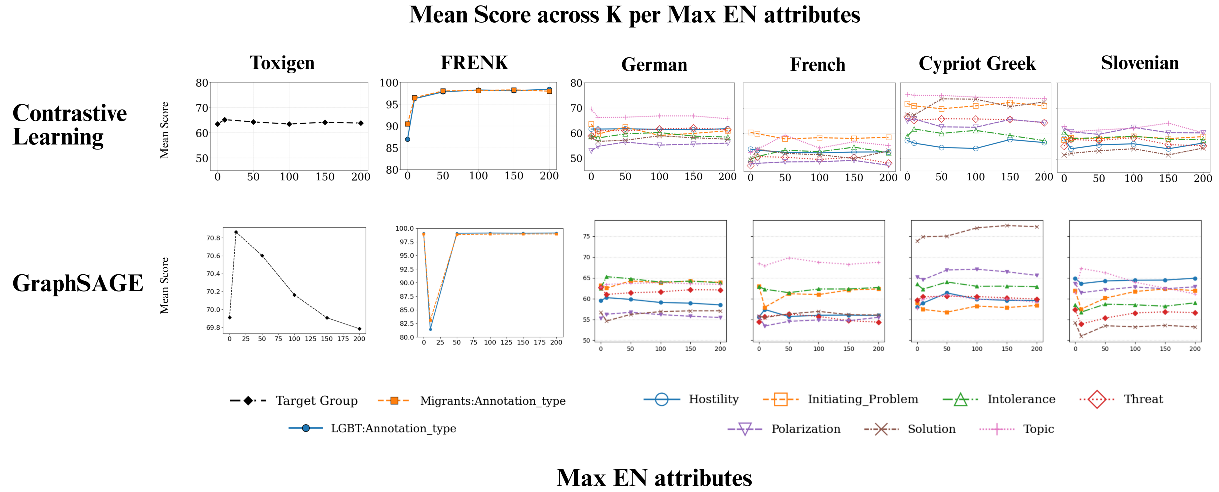

Figure 6: Node Classification in Toxigen, FRENK and the Multilingual EN Corpus: Mean Macro F1 scores for the target labels. Results are reported using node attributes, varying the **number of node features** considered per node (**Max EN attributes).**

representing gains of up to 15 points. FRENK-Migrants exhibits a comparable improvement, rising from 87–94 to 95–99 for Max EN $\geq$ 10. Performance is most stable for Max EN = 50–200 and k = 10–25, demonstrating that relational modeling with attribute-enriched nodes effectively enhances classification, particularly when a moderate to large number of auxiliary features are incorporated.

Multilingual EN Corpus In Tables 6, 7, 8, and 9, 10, 11, 12, 13) we report macro F1 scores for the Multilingual EN Corpus. Across all languages, incorporating node features consistently improves performance relative to the attribute-free baseline (Max EN = 0), whereas variations in neighborhood size (k) have limited impact.

For **German**, in **Semi-Supervised** graph learning, moderate feature sets (Max EN = 10–50) stabilize F1 scores, with gains of 2–3 points for labels such as *Intolerance* (62 $\rightarrow$ 65) and *Polarization* (55 $\rightarrow$ 57) compared with Max EN = 0, while variations in k have little effect. For **Contrastive Learning**, including node features increases F1 by 3–6 points across most labels, notably *Hostility, Intolerance, Threat, Polarization/Othering,* and *Solution*, with the largest gains observed when considering a small to moderate number of features (Max EN = 10–50), and limited additional improvement from larger feature sets.

For **French**, auxiliary features generally improve performance, though some minor drops occur in some settings. In **Semi-Supervised** learning, moderate Max EN (10–50) boosts F1 for labels like *Topic* (67 $\rightarrow$ 71.05), or *Threat* (54 $\rightarrow$ 58) while labels *Initiating Problem* show slight decreases at some settings (e.g., 63 $\rightarrow$ 58). Larger Max EN (100–200) stabilizes scores within 1–2 points. Larger improvements are observed for **Contrastive Learning**, with F1 gains of 4–14 points, particularly for *Topic, Threat, Hostility*, and *Polarization/Othering*. Performance stabilizes once a moderate number of node features is included, with additional features yielding only marginal gains.

In **Cypriot Greek**, in **Semi-Supervised** learning, moderate Max EN (10–50) improves labels like *Hostility* (58 $\rightarrow$ 62), *Intolerance* (62.74 $\rightarrow$ 64.55), and *Topic* (57 $\rightarrow$ 62) while some labels such as *Initiating Problem* show minor drops. Higher Max EN (100–200) stabilizes scores across all labels, with diminishing gains. For **Contrastive Learning** baseline performance is strong for *Topic* and *Initiating Problem* but lower for *Hostility* and *Intolerance*; incorporating node features substantially enhances these weaker categories, yielding gains up to 15 points before saturating at higher Max EN values.

In **Slovenian**, for **Semi-Supervised** learning, increasing node features and neighbors mainly improves *Topic*, while other labels such as *Hostility, Initiating Problem, Intolerance, Threat, Polarization,* and *Solution* show minor fluctuations or occasional drops. In **Contrastive Learning**, including node features improves weaker labels such as *Initiating Problem, Threat,* and *Solution* by 3 to 8 points, with peak performance at moderate Max EN values, while additional features provide smaller or

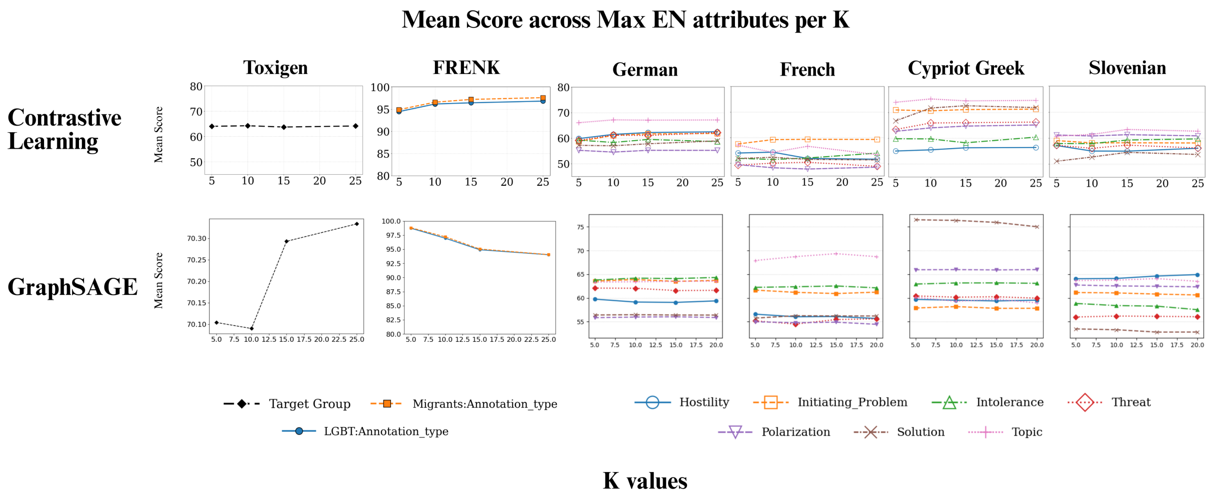

Figure 7: Node Classification in Toxigen, FRENK and the Multilingual EN Corpus: Mean Macro F1 scores for the target labels. Results are reported using node attributes, varying the **number of neighbors** (k) used to construct the graph.

less stable gains.

6.2 Overall Performance

Table 1 (Appendix) summarizes the results for models trained with and without attribute edges (Max EN $= 0$), reporting the minimum and maximum performance among the different tested settings. We compare our approach with a fine-tuned BERT Base uncased model (Devlin et al., 2019) for text classification using the [CLS] representation.

Across all datasets, incorporating node features consistently improves performance compared to models trained without them. Furthermore, models leveraging graph-based representations outperform the text-only BERT baseline, demonstrating the benefits of modeling relational and contextual dependencies in EN prediction tasks. Semi supervised learning with added EN attributes GraphSAGE yields stable improvements over both the baseline MLM and its KNN-only counterpart (referred to as *SS w/o node attr* in table 1). This suggests that propagating attribute information is beneficial beyond simple neighborhood-based supervision, with relative performance gains of up to +6.95% and a mean improvement of +1.90% across evaluated settings. However, in a small number of cases, when predicting EN attributes the KNN-only variant slightly outperforms the attribute-enriched model. This is the case for example of *Threat* in the German dataset and "Polarization" in the Slovenian dataset.

Finally, the contrastive learning approach provides a modest but consistent improvement over semi-supervised-only training. Incorporating explicit contrastive objectives results in relative gains of up to +10.86%, with a mean improvement of +2.55%, suggesting that contrastive supervision further enhances the model's ability to capture structural relationships in the graph.

7 Conclusion

We presented GR-EN-A-DE, a novel graph-based Extremist Narrative model upon which we proposed a classification framework based on two graph learning strategies: a semi-supervised encoder–decoder and a contrastive learning approach. Experiments on different available and newly proposed datasets show that incorporating relational and node-level information consistently improves performance over text-only baselines, emphasizing the importance of modeling structural and contextual dependencies. These findings demonstrate the potential of graph-based methods for analyzing complex social and linguistic phenomena and motivate future exploration of graph architectures and feature integration in multilingual and cross-domain settings.

Limitations

Extremist narratives are inherently complex and highly context-dependent, making annotation challenging and introducing subjectivity into the labels. Annotator confidence in the Multilingual EN Corpus was generally moderate, underscoring the difficulty of consistently identifying and categorizing extremist content. These factors may limit the reliability of the training data and, in turn, affect model performance. The results of our effort, which is among the first in this direction, indicate that leveraging relational and contextual information through graph-based methods is a more effective approach for modeling and classifying extremist content. To extend our work, ad-hoc graph-based explainability is essential to understand how and under which conditions the adopted models capture and leverage complex extremist narrative structures such as actors and actants (e.g., in-group versus out-group), role configurations (victim, aggressor, defender), patterns of conflict, causal links, and underlying moral or ideological frames.

References

Gordon Willard Allport, Kenneth Clark, and Thomas F Pettigrew. 1954. *The nature of prejudice*, volume 2. Addison-wesley Reading, MA.

Selene Baez Santamaria, Helena Gomez Adorno, and Ilia Markov. 2024. Contextualized graph representations for generating counter-narratives against hate speech. In *Findings of the Association for Computational Linguistics: EMNLP 2024*, pages 7664–7674, Miami, Florida, USA. Association for Computational Linguistics.

Fabienne Baider and Alexandras Gregoriou. 2024. Annotating linguistic features of extreme narratives: a literature review and a proposal. *Studii de Lingvistică*, 14(2).

Fabienne Baider and Monika Kopytowska. 2018. Narrating hostility, challenging hostile narratives. *Lodz Papers in Pragmatics*, 14(1):1–24.

Matthew Beatty. 2020. Graph-based methods to detect hate speech diffusion on twitter. In *2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, pages 502–506.

Heiko Beyer and Alexandra Schauer. 2021. *Die Rückkehr der Ideologie. Zur Gegenwart eines Schlüsselbegriffs*.

Necva Bölücü and Pelin Canbay. 2021. Hate speech and offensive content identification with graph convolutional networks. In *Working Notes of FIRE 2021 - Forum for Information Retrieval Evaluation, Gandhinagar, India, December 13-17, 2021*, volume 3159 of *CEUR Workshop Proceedings*, pages 44–51. CEUR-WS.org.

Katy Brown, Aurelien Mondon, and Aaron Winter. 2023. The far right, the mainstream and mainstreaming: towards a heuristic framework. *Journal of Political Ideologies*, 28(2):162–179.

Souvic Chakraborty, Parag Dutta, Sumegh Roychowdhury, and Animesh Mukherjee. 2022. Crush: Contextually regularized and user anchored self-supervised hate speech detection. *Preprint*, arXiv:2204.06389.

David Chavalarias, Paul Bouchaud, and Maziyar Panahi. 2023. Can few lines of code change society ? beyond fack-checking and moderation : how recommender systems toxifies social networking sites. *Preprint*, arXiv:2303.15035.

Thomas M Cover and Peter E Hart. 1967. Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1):21–27.

Marco Del Tredici, Diego Marcheggiani, Sabine Schulte im Walde, and Raquel Fernández. 2019. You shall know a user by the company it keeps: Dynamic representations for social media users in NLP. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 4707–4717, Hong Kong, China. Association for Computational Linguistics.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. Bert: Pre-training of deep bidirectional transformers for language understanding. *Preprint*, arXiv:1810.04805.

Charles Duong, Lei Zhang, and Chang-Tien Lu. 2022. Hatenet: A graph convolutional network approach to hate speech detection. In *2022 IEEE International Conference on Big Data (Big Data)*, pages 5698–5707.

Robert M Fano. 1963. *Transmission of information: a statistical theory of communications*. MIT press.

Steven Forti. 2021. *Extrema derecha 2.0.: qué es y cómo combatirla*. Siglo XXI de España Editores.

Noé Gaumont, Maziyar Panahi, and David Chavalarias. 2018. Reconstruction of the socio-semantic dynamics of political activist twitter networks—method and application to the 2017 french presidential election. *PloS one*, 13(9):e0201879.

Andrew Glazzard. 2017. Losing the plot: Narrative, counter-narrative and violent extremism. *Terrorism and Counter-Terrorism Studies*, 8.

William L. Hamilton, Rex Ying, and Jure Leskovec. 2018. Inductive representation learning on large graphs. *Preprint*, arXiv:1706.02216.

Thomas Hartvigsen, Saadia Gabriel, Hamid Palangi, Maarten Sap, Dipankar Ray, and Ece Kamar. 2022. Toxigen: A large-scale machine-generated dataset for adversarial and implicit hate speech detection. *Preprint*, arXiv:2203.09509.

Liam Hebert, Hong Yi Chen, Robin Cohen, and Lukasz Golab. 2023. Qualitative analysis of a graph transformer approach to addressing hate speech: Adapting to dynamically changing content. *Preprint*, arXiv:2301.10871.

Liam Hebert, Lukasz Golab, and Robin Cohen. 2022. Predicting hateful discussions on reddit using graph transformer networks and communal context. In *2022 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)*, pages 9–17.

Liam Hebert, Gaurav Sahu, Yuxuan Guo, Nanda Kishore Sreenivas, Lukasz Golab, and Robin Cohen. 2024. Multi-modal discussion transformer: Integrating text, images and graph transformers to detect hate speech on social media. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pages 22096–22104.

Oliver Hidalgo. 2021. (rechts-)populismus als ideologie? Überlegungen zu einem wiederkehrenden strukturdefekt der demokratie. In H. Beyer and A. Schauer, editors, *Die Rückkehr der Ideologie: Zur Gegenwart eines Schlüsselbegriffs*, pages 369–392. Campus.

Haroro Ingram. 2016a. A "linkage-based" approach to combating militant islamist propaganda: A two-tiered framework for practitioners. *Terrorism and Counter-Terrorism Studies*, 7.

Haroro J Ingram. 2016b. A "linkage-based" approach to combating militant islamist propaganda: A two-tiered framework for.

Gilles Ivaldi. 2024. The populist radical-right turn of the mainstream right in france. *The transformation of the mainstream right and its impact on (social) democracy*, pages 80–93.

Wenhao Jiang and Yuebin Bai. 2024. Sgcl: Semi-supervised graph contrastive learning with confidence propagation algorithm for node classification. *Knowledge-Based Systems*, 301:112271.

Thorsten Joachims and 1 others. 1997. A probabilistic analysis of the rocchio algorithm with tfidf for text categorization. In *ICML*, volume 97, pages 143–151. Citeseer.

Aristotle Kallis. 2013. Far-right "contagion" or a failing "mainstream"? how dangerous ideas cross borders and blur boundaries. *Democracy and Security*, 9(3):221–246.

Yixin Liu, Yu Zheng, Daokun Zhang, Hongxu Chen, Hao Peng, and Shirui Pan. 2022. Towards unsupervised deep graph structure learning. *Preprint*, arXiv:2201.06367.

Nikola Ljubešić, Darja Fišer, and Tomaž Erjavec. 2019. The frenk datasets of socially unacceptable discourse in slovene and english. *Preprint*, arXiv:1906.02045.

Simo K Määttä. 2023. Linguistic and discursive properties of hate speech and speech facilitating the expression of hatred: evidence from finnish and french online discussion boards. *Internet Pragmatics*, 6(2):156–172.

Pushkar Mishra, Marco Del Tredici, Helen Yannakoudakis, and Ekaterina Shutova. 2018. Author profiling for abuse detection. In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 1088–1098, Santa Fe, New Mexico, USA. Association for Computational Linguistics.

Pushkar Mishra, Marco Del Tredici, Helen Yannakoudakis, and Ekaterina Shutova. 2019. Abusive Language Detection with Graph Convolutional Networks. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 2145–2150, Minneapolis, Minnesota. Association for Computational Linguistics.

Xinyi Mou, Zhongyu Wei, Lei Chen, Shangyi Ning, Yancheng He, Changjian Jiang, and Xuanjing Huang. 2021. Align voting behavior with public statements for legislator representation learning. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 1236–1246, Online. Association for Computational Linguistics.

Yufei Mu, Jin Yang, Tianrui Li, Siyu Li, and Weiheng Liang. 2024. Ha-gcen: Hyperedge-abundant graph convolutional enhanced network for hate speech detection. *Knowledge-Based Systems*, 300:112166.

Cas Mudde. 2007. *Populist Radical Right Parties in Europe*.

Cas Mudde and Cristóbal Rovira Kaltwasser. 2017. *Populism: A Very Short Introduction*. Oxford University Press.

Andreas Musolff. 2015. Dehumanizing metaphors in uk immigrant debates in press and online media. *Journal of Language aggression and conflict*, 3(1):41–56.

Curtis G. Northcutt, Lu Jiang, and Isaac L. Chuang. 2022. Confident learning: Estimating uncertainty in dataset labels. *Preprint*, arXiv:1911.00068.

A. Y. Postigo-Fuentes, S. Soler, and K. Pahor de Maiti Tekavčič. 2025. An interpretative framework for operationalization of extremist narrative analysis.

Ana Y. Postigo-Fuentes, Sara Soler, and Katarina Pahor de Maiti Tekavčič. in press. Operationalizing extremist narrative analysis: An interpretative framework. *Politics, Religion & Ideology*. Forthcoming.

Ana Yara Postigo Fuentes, Stefan Hartmann, Rolf Kailuweit, and Alexander Ziem. 2024. Defining extremist narratives: A review of the current state of the art.

Alastair Reed and Jennifer Dowling. 2018. The role of historical narratives in extremist propaganda. *Defence Strategic Communications*, 4(1):79–113.

Nils Reimers and Iryna Gurevych. 2019. Sentence-bert: Sentence embeddings using siamese bert-networks. *Preprint*, arXiv:1908.10084.

John E. Richardson and Ruth Wodak. 2022. Anti-sorosism: Reviving the "jewish world conspiracy". In *Discourse Approaches to Politics, Society and Culture*, pages 395–420. John Benjamins Publishing Company.

Kevin Smets, Koen Leurs, Myria Georgiou, Saskia Witteborn, and Radhika Gajjala, editors. 2019. *The SAGE Handbook of Media and Migration*. SAGE Publications Ltd, London.

Sahana Udupa, Iginio Gagliardone, and Peter Hervik. 2021. *Digital hate: The global conjuncture of extreme speech*. Indiana University Press.

Teun A Van Dijk and 1 others. 1998. Opinions and ideologies in the press. *Approaches to media discourse*, 21(63):1–33.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. 2023. Attention is all you need. *Preprint*, arXiv:1706.03762.

Francesco Veri and Franziska Maier. 2025. Migration discourses from the radical right: Mapping and testing potential for political mobilization. *Political Psychology*. Epub ahead of print.

Azmine Toushik Wasi. 2023. Explainable identification of hate speech towards islam using graph neural networks. *arXiv preprint arXiv:2311.04916*.

Kun Wu, Xinyi Mou, Lanqing Xue, Zhenzhe Ying, Weiqiang Wang, Qi Zhang, Xuanjing Huang, and Zhongyu Wei. 2024. PASUM: A pre-training architecture for social media user modeling based on text graph. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 12644–12656, Torino, Italia. ELRA and ICCL.

Keyulu Xu, Chengtao Li, Yonglong Tian, Tomohiro Sonobe, Ken ichi Kawarabayashi, and Stefanie Jegelka. 2018. Representation learning on graphs with jumping knowledge networks. *Preprint*, arXiv:1806.03536.

Yuning You, Tianlong Chen, Yongduo Sui, Ting Chen, Zhangyang Wang, and Yang Shen. 2020. Graph contrastive learning with augmentations. *ArXiv*, abs/2010.13902.

A Appendix

A.1 Benchmark Datasets

A.1.1 ToxiGen

Dataset Overview. ToxiGen is an open-source, large-scale benchmark English dataset developed by Hartvigsen et al. (2022) for detecting and evaluating toxic and socially biased language. It captures both explicit and implicit expressions of toxicity across a broad spectrum of social groups and identity categories, addressing the limitations of earlier resources that primarily focused on overt hate speech or profanity. The dataset contains 274K statements-approximately half toxic and half benign-generated through a combination of human-authored prompts and machine-generated continuations using the GPT-3 language model. The data generation process employed demonstration-based prompting, a technique in which example sentences guide the model to produce new statements that mention minority groups in both neutral and biased ways. ToxiGen spans 13 minority identity groups, including racial, gender, sexual, religious, and disability-related categories (e.g., African American, Muslim, LGBTQ+, Women, Physically or Mentally Disabled, Jewish, Asian, and Latino). The resulting corpus consists largely of human-like text, most of which lacks explicit slurs, making it suitable for studying implicit prejudice and model bias.

Labels and Auxiliary Attributes. Each instance in ToxiGen is accompanied by multiple annotations describing how social groups are referenced and discussed. In our setting, the *target_group* field is used to identify the social or identity group mentioned in the text and serves as the primary signal for associating instances with targets.

Additional annotations provide complementary information about the linguistic and pragmatic properties of the statement. Specifically, the *ingroup_effect* field indicates whether identity-aligned or self-referential language is used, while *framing* characterizes the communicative intent of the statement (e.g., moral judgment, disagreement, or causal explanation). These annotations are incorporated as auxiliary attributes to enrich node representations.

Data Filtering. To ensure sufficient representation for model training, we excluded minority classes that individually accounted for less than 5% of the data. These removals correspond to a small fraction of the total dataset, ensuring that the overall coverage and distribution remain largely unaffected.

A.1.2 FRENK

Dataset Overview. FRENK (Ljubešić et al., 2019) is a collection of Facebook comment threads responding to mainstream media posts in Slovene and English. Developed within the Slovene national project FRENK, the dataset focuses on two topics (migrants and LGBT) and is manually annotated for both the type of socially unacceptable discourse (SUD) and its targets. The annotation schema distinguishes six types of SUD (*acceptable speech*, *inappropriate speech*, *background offensive*, *background violence*, *other offensive*, *other threat/violence*) and five possible targets (*migrants or LGBT individuals*, *related groups or organizations*, *journalists and media*, *other commenters*, *others*).

The dataset comprises 30 Slovene posts on migrants, 93 on LGBT, 16 English posts on migrants, and 14 on LGBT, totaling roughly 4,500-6,500 comments per topic. Each comment received multiple annotations, averaging 7-9 per comment, resulting in over 40K annotations per topic and language.

Labels and Auxiliary Attributes. We use the *annotation_type* field as the primary label, indicating the type of socially unacceptable speech. Complementary information is captured through the anonymized *user_id* and the *annotation_target* fields. The *user_id* reflects the commenter's perspective, implicitly providing an in-group signal, while *annotation_target* identifies the referenced group or entity similarly to the out-group. These fields facilitate the construction of relational edges and node attributes in the graph, supporting the modeling of interaction and reference patterns without exposing personal information.

Data Filtering. For our experiments, we used only the English subset of FRENK and aggregated multiple annotations per comment via majority vote. Similar to ToxiGen, rare classes representing less than 5% of the data were excluded. These preprocessing steps retain the majority of the dataset while ensuring reliable and consistent labels for model training.

A.1.3 Multilingual EN Corpus

Data Selection and Collection. The selection and collection of data followed a structured and collaborative process between the research teams and the data engineers to ensure alignment between the project's research goals and technical feasibility. Preliminary discussions led to identifying the relevant digital platforms (Twitter, Facebook, TikTok, YouTube, Telegram, and Wikipedia) as key arenas of political discourse across Europe. Researchers in each country mapped social media accounts of politicians, political parties, institutions, and other influential actors deemed relevant for the analysis. The inclusion criteria focused on accounts that produced or shared content aligned with Extremist Narratives, with particular attention to right-wing ideologies. Content was selected if it addressed topics such as science, gender, or nation in ways that reflected closed, non-negotiable, or potentially violent discourse, or that displayed chauvinistic, xenophobic, misogynistic, or homophobic motives, or a rejection of scientific knowledge. This focus is theoretically justified by the growing political and social impact of far-right ideologies across Europe, which often stand in opposition to the core democratic values of inclusivity, equality, and mutual respect that underpin the European Union (Mudde, 2007; Forti, 2021). As noted in recent studies (Kallis, 2013; Brown et al., 2023; Ivaldi, 2024), far-right ideas are increasingly adopted or normalized within mainstream right-wing parties, further motivating a broader analysis of right-wing discourse online. Data collection was conducted through platform APIs (Twitter, TikTok, YouTube, and Wikipedia) while data from Facebook and Telegram were limited to publicly available content due to technical restrictions. The process ensured that both qualitative and quantitative data were collected in a transparent, systematic manner, including textual content, nodes, and engagement metrics where available.

Annotation Process. A subset of the collected dataset was subjected to manual annotation to enable systematic qualitative analysis. In total, 1,002 tweets (334 per topic) were annotated by two independent annotators per language following the annotation schema developed and published in Baider and Gregoriou (2024). This schema provided the theoretical and methodological framework for identifying and classifying relevant discursive elements in the corpus, ensuring that the annotation process was consistent and replicable across annotators.

Annotation Quality Assessment. Each dataset was independently annotated by two expert annotators. To evaluate annotation reliability, we applied the Cleanlab framework (Northcutt et al., 2022),

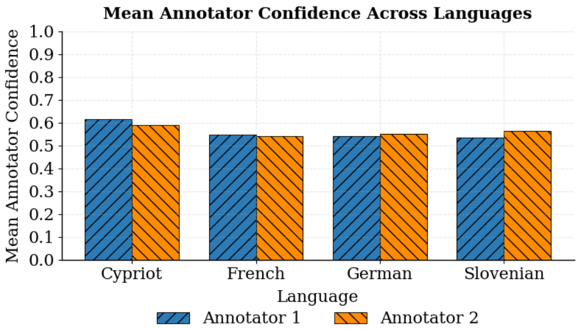

Figure 8: Mean annotator confidence across languages for target features: Perceived Threat, Initiating Problem, Polarization/Othering, Hostility to out-group.

which estimates the probability of label correctness using confident learning. Sentence embeddings were generated with a SentenceTransformer model (MiniLM-L3-v2) to represent each instance in a high-dimensional semantic space. A simple logistic regression classifier with cross-validation was then used to predict label probabilities for each feature. For each language, and class, we computed the mean Cleanlab confidence score per annotator, representing the expected correctness of labels. Mean annotator confidence ranged between 0.5 to 0.6 across languages reflecting the task's inherent complexity and thus the moderate certainty for target features(Perceived Threat, Initiating Problem, Polarization/Othering, Hostility to out-group). On the other hand, we note that Confidence is very similar across annotators.

Annotation Alignment To derive a single gold label from two independent annotations, we applied a deterministic set of disagreement-resolution rules designed to favor unambiguous interpretations. Prior to merging, duplicate texts were removed from the dataset, retaining only the first occurrence to prevent redundancy and label leakage.

For binary labels (True/False), any disagreement was resolved by assigning the negative class, reflecting a conservative bias against over-annotation. For ordinal or scalar labels, disagreements were resolved by selecting the lower or weaker value. Specifically, *Tone of Post* was set to the lower-valued tone, while *Topic* and *Narrator* were set to *Irrelevant* and *Unclear*, respectively, in cases of disagreement.

Disagreements for In-Group and Out-Group labels were resolved using a hierarchy based on political orientation and specificity. If both annotators selected values within the same political spectrum but differed in extremity (e.g., *Left* vs. *Far-Left*), the less extreme category was selected. If annotators selected opposing political orientations (e.g., *Left/Far-Left* vs. *Right/Far-Right*), the label was set to *Political Party(ies)*, indicating a generalized political reference. When one annotation was *Other Country* and the other *Other*, the merged label was set to *Other*. All remaining unresolved disagreements were assigned *None/Unclear*.

If a label was missing from one annotation but present in the other, the present label was retained, assuming annotator omission rather than disagreement. For the *Character(s)* attribute, when one annotation was *Explicit* and the other *Implicit*, the merged label was set to *Implicit*, reflecting a more cautious interpretation.

A.2 Detailed Experimental Results

Dataset	Method	Target Group	Annot. Type	Hostility	Init. Problem	Intolerance	Threat	Polarization	Solution	Topic
Toxigen	SS w/o node attr.	-	-	-	-	-	-	-	-	-
	SS w node attr.	71.80$_{0.372}$	-	-	-	-	-	-	-	-
	CL w/o node attr.	63.69$_{0.85}$	-	-	-	-	-	-	-	-
	CL w node attr.	65.60$_{1.21}$	-	-	-	-	-	-	-	-
	Baseline	71.27$_{1.21}$	-	-	-	-	-	-	-	-
LGBT	SS w/o node attr	-	99.27$_{0.12}$	-	-	-	-	-	-	-
	SS w/ node attr	-	99.19$_{0.06}$	-	-	-	-	-	-	-
	CL w/o node attr	-	89.870.80	-	-	-	-	-	-	-
	CL w/ node attr	-	98.780.35	-	-	-	-	-	-	-
	Baseline	-	65.12$_{3.00}$	-	-	-	-	-	-	-
Migrants	SS w/o node attr	-	98.950.00	-	-	-	-	-	-	-
	SS w/ node attr	-	98.950.00	-	-	-	-	-	-	-
	CL w/o node attr	-	93.950.40	-	-	-	-	-	-	-
	CL w/ node attr	-	98.880.07	-	-	-	-	-	-	-
	Baseline	-	67.24$_{0.78}$	-	-	-	-	-	-	-
GER	SS w/o node attr	-	-	(58.71-61.28)$_{1.18}$	(62.70-63.63)$_{0.40}$	(62.20-62.96)$_{0.33}$	(61.91-63.20)$_{0.58}$	(54.81-56.07)$_{0.58}$	(56.03-57.77)$_{0.75}$	(62.33-63.29)$_{0.42}$
	SS w/ node attr	-	-	(58.03-60.86)$_{0.80}$	(62.04-64.48)$_{0.71}$	(63.34-65.46)$_{0.68}$	(60.19-62.62)$_{0.61}$	(54.79-57.18)$_{0.69}$	(54.02-57.97)$_{1.08}$	(62.74-64.20)$_{0.35}$
	CL w/o node attr	-	-	(59.55-63.59)$_{1.95}$	(62.78-65.22)$_{1.15}$	(55.96-61.44)$_{2.53}$	(57.90-59.88)$_{0.98}$	(50.53-54.36)$_{1.65}$	(56.79-61.43)$_{2.12}$	(68.23-70.18)$_{0.89}$
	CL w/ node attr	-	-	(58.61-63.71)$_{1.46}$	(54.32-63.52)$_{2.56}$	(55.62-61.89)$_{1.63}$	(58.16-63.47)$_{1.50}$	(52.04-58.38)$_{1.57}$	(52.98-59.60)$_{1.79}$	(63.64-67.96)$_{1.12}$
	Baseline	-	-	44.06 $_{4.29}$	52.77 $_{8.24}$	57.03 $_{3.05}$	50.91 $_{4.53}$	47.81 $_{6.61}$	51.72 $_{7.73}$	60.16 $_{3.45}$
FR	SS w/o node attr	-	-	(54.93-56.26)$_{0.55}$	(62.05-63.46)$_{0.65}$	(62.61-63.16)$_{0.24}$	(54.28-54.49)$_{0.09}$	(55.34-56.38)$_{0.49}$	(55.43-56.02)$_{0.28}$	(67.43-69.33)$_{0.89}$
	SS w/ node attr	-	-	(54.95-58.22)$_{0.88}$	(57.67-63.03)$_{1.66}$	(60.60-63.14)$_{0.65}$	(53.64-58.00)$_{1.24}$	(52.75-56.04)$_{0.85}$	(54.29-57.55)$_{0.70}$	(66.75-71.05)$_{0.97}$
	CL w/o node attr	-	-	(51.14-56.22)$_{2.18}$	(58.21-62.36)$_{1.73}$	(48.07-52.52)$_{1.88}$	(47.14-47.80)$_{0.30}$	(46.67-49.38)$_{1.41}$	(48.21-51.29)$_{1.29}$	(50.92-56.47)$_{2.64}$
	CL w/ node attr	-	-	(47.66-57.02)$_{2.86}$	(53.41-62.69)$_{2.20}$	(49.06-57.48)$_{2.49}$	(47.52-52.99)$_{1.81}$	(45.84-51.43)$_{1.50}$	(48.86-57.05)$_{2.27}$	(51.22-65.01)$_{3.88}$
	Baseline	-	-	48.91 $_{2.79}$	54.36 $_{6.44}$	52.60 $_{6.18}$	48.35 $_{0.22}$	48.74 $_{2.58}$	48.73 $_{0.27}$	51.2 $_{3.70}$
CY	SS w/o node attr	-	-	(57.52-58.14)$_{0.29}$	(58.21-59.50)$_{0.55}$	(62.74-64.17)$_{0.70}$	(59.31-59.97)$_{0.30}$	(64.78-65.44)$_{0.28}$	(73.12-75.06)$_{0.87}$	(56.61-58.56)$_{0.84}$
	SS w/ node attr	-	-	(58.54-62.18)$_{0.92}$	(55.73-58.75)$_{0.74}$	(60.56-64.55)$_{0.92}$	(58.86-61.87)$_{0.64}$	(64.02-67.83)$_{1.09}$	(73.37-78.90)$_{1.59}$	(58.37-61.67)$_{0.83}$
	CL w/o node attr	-	-	(56.12-58.50)$_{1.12}$	(70.94-72.18)$_{0.55}$	(56.61-62.25)$_{2.44}$	(65.35-68.05)$_{1.31}$	(63.82-65.46)$_{0.69}$	(62.29-74.49)$_{5.64}$	(74.91-75.86)$_{0.41}$
	CL w/ node attr	-	-	(51.72-60.48)$_{2.16}$	(68.75-73.17)$_{1.27}$	(55.17-64.55)$_{2.69}$	(61.73-68.64)$_{1.83}$	(59.10-67.14)$_{2.17}$	(62.86-76.77)$_{4.19}$	(71.78-76.15)$_{1.14}$
	Baseline	-	-	44.86 $_{0.54}$	45.63 $_{6.25}$	46.35 $_{0.81}$	44.21 $_{0.37}$	50.38 $_{5.10}$	48.51 $_{0.09}$	38.51 $_{9.20}$
SL	SS w/o node attr	-	-	(64.51-65.11)$_{0.26}$	(61.50-62.14)$_{0.28}$	(58.06-58.88)$_{0.34}$	(56.27-57.75)$_{0.70}$	(62.71-64.67)$_{0.82}$	(53.72-54.65)$_{0.46}$	(60.64-62.19)$_{0.74}$
	SS w/ node attr	-	-	(63.17-65.35)$_{0.66}$	(55.83-62.70)$_{1.93}$	(55.54-59.51)$_{1.04}$	(52.88-57.71)$_{1.26}$	(60.48-63.32)$_{0.69}$	(50.46-54.02)$_{1.08}$	(60.81-68.06)$_{2.41}$
	CL w/o node attr	-	-	(56.78-60.29)$_{1.54}$	(54.59-60.38)$_{2.44}$	(58.99-63.27)$_{1.85}$	(53.04-58.69)$_{2.40}$	(60.44-66.22)$_{2.74}$	(49.48-54.65)$_{2.58}$	(61.10-65.69)$_{2.19}$
	CL w/ node attr	-	-	(51.78-59.54)$_{2.16}$	(56.32-60.77)$_{1.26}$	(56.35-60.50)$_{1.26}$	(53.13-60.53)$_{1.80}$	(57.99-64.08)$_{1.56}$	(47.73-56.33)$_{2.01}$	(58.08-65.46)$_{1.96}$
	Baseline	-	-	43.43 $_{4.09}$	49.02 ± $_{5.20}$	48.68 ± $_{5.01}$	45.50 $_{5.75}$	45.67 $_{3.21}$	46.02 $_{0.68}$	49.15 $_{4.01}$

Table 1: F1 score comparison of semi-supervised (SS) and contrastive learning (CL) node classification models, with and without node feature representations, and a BERT-base text classification baseline, across datasets (mean $_{SD}$).

Max	K=5	K=10	K=15	K=25
0	68.370$_{0.300}$	69.486$_{1.087}$	70.905$_{0.793}$	70.884$_{0.671}$
10	71.801$_{0.372}$	70.542$_{0.190}$	70.387$_{0.334}$	70.723$_{0.290}$
50	70.602$_{0.289}$	70.641$_{0.311}$	70.581$_{0.281}$	70.581$_{0.471}$
100	70.152$_{0.362}$	70.160$_{0.321}$	70.186$_{0.326}$	70.154$_{0.399}$
150	69.898$_{0.445}$	69.915$_{0.389}$	69.888$_{0.355}$	69.927$_{0.417}$
200	69.801$_{0.334}$	69.796$_{0.244}$	69.811$_{0.359}$	69.732$_{0.272}$

Table 2: **Semi-supervised (GraphSAGE)** Node Classification in **Toxigen**: Macro F1 scores for the target label *target_group*. Results are reported using *ingroup_effect* and *framing* node attributes, varying the **number of node features** considered per node (Max EN) and the **number of neighbors** (k) used to construct the graph. Minority classes (less than 5% of dataset elements) were removed.

Max	K=5	K=10	K=15	K=25
0	63.72$_{0.76}$	63.69$_{0.85}$	62.88$_{1.10}$	63.48$_{0.74}$
10	**64.89**$_{0.52}$	**65.55**$_{0.66}$	**64.67**$_{1.04}$	**65.60**$_{0.20}$
50	63.97$_{0.72}$	64.11$_{0.96}$	64.13$_{0.52}$	64.88$_{0.66}$
100	63.53$_{0.68}$	63.97$_{1.25}$	62.93$_{0.52}$	63.32$_{0.32}$
150	64.27$_{0.80}$	63.78$_{1.02}$	64.37$_{0.48}$	64.10$_{0.97}$
200	63.75$_{0.70}$	64.44$_{0.87}$	63.50$_{0.65}$	63.53$_{0.20}$

Table 3: **Contrastive Learning** Node Classification in **Toxigen**: Macro F1 scores for the target label *target_group*. Results are reported using *ingroup_effect* and *framing* node attributes, varying the **number of node features** considered per node (Max EN) and the **number of neighbors** (k) used to construct the graph. Minority classes (less than 5% of dataset elements) were removed.

Max	LGBT				Migrants			
	K=5	K=10	K=15	K=25	K=5	K=10	K=15	K=25
0	98.62$_{0.00}$	99.17$_{0.12}$	99.27$_{0.12}$	99.27$_{0.12}$	98.95$_{0.00}$	98.95$_{0.00}$	98.95$_{0.00}$	98.95$_{0.00}$
10	97.48$_{0.09}$	86.33$_{0.94}$	73.59$_{0.89}$	68.37$_{0.56}$	98.16$_{0.11}$	88.65$_{0.31}$	75.75$_{0.26}$	69.74$_{0.56}$
50	99.01$_{0.12}$	99.06$_{0.14}$	99.17$_{0.12}$	99.11$_{0.14}$	98.90$_{0.07}$	98.84$_{0.14}$	98.77$_{0.17}$	98.87$_{0.12}$
100	99.11$_{0.14}$	99.06$_{0.14}$	99.13$_{0.12}$	99.19$_{0.06}$	98.92$_{0.06}$	98.92$_{0.06}$	98.95$_{0.00}$	98.92$_{0.06}$
150	99.06$_{0.14}$	99.06$_{0.14}$	99.17$_{0.12}$	99.11$_{0.14}$	99.00$_{0.07}$	98.92$_{0.06}$	98.92$_{0.06}$	98.95$_{0.00}$
200	99.11$_{0.14}$	99.06$_{0.14}$	99.17$_{0.12}$	99.17$_{0.12}$	98.97$_{0.06}$	98.95$_{0.00}$	98.95$_{0.00}$	98.95$_{0.00}$

Table 4: **Semi-supervised (GraphSAGE based)** Node Classification in **FRENK**: Macro F1 scores for the target label *annotation_type*. Results are reported using *user_id* and *annotation_target* node attributes, varying the **number of node features** considered per node (Max EN) and the **number of neighbors** (k) used to construct the graph. Minority classes (less than 5% of dataset elements) were removed.

Max	LGBT				Migrants			
	K=5	K=10	K=15	K=25	K=5	K=10	K=15	K=25
0	$83.17_{0.51}$	$86.59_{0.96}$	$88.07_{0.65}$	$89.87_{0.80}$	$86.88_{0.41}$	$89.61_{0.40}$	$91.42_{0.28}$	$93.95_{0.40}$
10	$94.70_{0.38}$	$96.76_{0.16}$	$96.87_{0.51}$	$96.65_{0.31}$	$95.02_{0.22}$	$96.60_{0.18}$	$97.25_{0.28}$	$97.04_{0.28}$
50	$96.84_{0.41}$	$98.00_{0.17}$	$98.20_{0.40}$	$98.15_{0.53}$	$96.06_{0.42}$	$98.30_{0.18}$	$\mathbf{98.88}_{0.07}$	$\mathbf{98.81}_{0.18}$
100	$97.26_{0.34}$	$98.66_{0.33}$	$98.39_{0.13}$	$98.65_{0.19}$	$\mathbf{97.51}_{0.29}$	$98.20_{0.11}$	$98.37_{0.07}$	$98.38_{0.27}$
150	$96.97_{0.26}$	$98.08_{0.15}$	$98.40_{0.25}$	$\mathbf{98.77}_{0.10}$	$97.15_{0.21}$	$\mathbf{98.53}_{0.07}$	$98.67_{0.10}$	$98.68_{0.16}$
200	$\mathbf{97.63}_{0.40}$	$\mathbf{98.78}_{0.35}$	$\mathbf{98.57}_{0.24}$	$98.70_{0.32}$	$96.63_{0.42}$	$98.08_{0.24}$	$98.49_{0.13}$	$98.64_{0.23}$

Table 5: **Contrastive Learning** Node Classification in **FRENK**: Macro F1 scores for the target label *annotation_type*. Results are reported using *user_id* and *annotation_target* node attributes, varying the **number of node features** considered per node (Max EN) and the **number of neighbors** (k) used to construct the graph. Minority classes (less than 5% of dataset elements) were removed.

Max	K	Hostility	Initiating Problem	Intolerance	Threat	Polarization	Solution	Topic
0	5	$61.28_{1.81}$	$62.70_{1.76}$	$62.35_{1.66}$	$63.20_{0.51}$	$54.90_{1.21}$	$56.03_{0.36}$	$62.33_{0.91}$
	10	$58.71_{1.38}$	$63.63_{2.17}$	$62.96_{1.12}$	$62.48_{1.62}$	$55.41_{1.06}$	$56.38_{0.96}$	$62.55_{0.47}$
	15	$59.09_{1.76}$	$63.07_{2.09}$	$62.20_{1.40}$	$63.01_{1.72}$	$56.07_{1.16}$	$57.77_{0.41}$	$62.88_{0.85}$
	25	$59.04_{0.90}$	$63.36_{1.46}$	$62.39_{0.82}$	$61.91_{0.73}$	$54.81_{1.62}$	$56.80_{1.19}$	$63.29_{0.80}$
10	5	$59.82_{2.67}$	$62.04_{2.01}$	$65.40_{1.25}$	$61.16_{1.06}$	$56.92_{1.39}$	$54.74_{1.25}$	$63.55_{0.99}$
	10	$60.86_{1.60}$	$62.54_{1.22}$	$65.05_{0.93}$	$61.73_{2.44}$	$56.93_{2.10}$	$55.07_{0.92}$	$63.48_{0.98}$
	15	$59.37_{1.39}$	$62.60_{1.55}$	$65.11_{0.55}$	$60.19_{1.68}$	$56.34_{1.99}$	$54.02_{1.31}$	$63.87_{0.90}$
	25	$60.84_{1.85}$	$62.92_{2.32}$	$65.46_{1.21}$	$60.90_{1.37}$	$54.79_{1.63}$	$54.67_{1.18}$	$62.74_{1.22}$
50	5	$60.72_{1.72}$	$63.95_{1.12}$	$63.72_{2.55}$	$60.94_{1.22}$	$57.18_{0.75}$	$56.32_{1.42}$	$63.95_{1.29}$
	10	$59.32_{0.66}$	$64.45_{1.23}$	$64.81_{1.36}$	$62.07_{1.91}$	$56.08_{1.72}$	$57.00_{0.74}$	$63.23_{0.82}$
	15	$59.13_{1.11}$	$64.08_{1.09}$	$65.31_{1.89}$	$60.95_{1.32}$	$56.38_{0.82}$	$56.19_{1.88}$	$63.33_{0.48}$
	25	$60.04_{1.75}$	$64.48_{1.26}$	$65.24_{2.24}$	$61.70_{0.70}$	$57.18_{1.73}$	$55.27_{1.51}$	$64.12_{1.19}$
100	5	$58.92_{1.00}$	$63.93_{1.03}$	$63.34_{2.58}$	$62.34_{2.26}$	$55.75_{1.16}$	$57.77_{1.12}$	$63.36_{0.92}$
	10	$59.14_{1.59}$	$64.22_{0.71}$	$64.11_{1.35}$	$61.72_{1.59}$	$55.90_{1.69}$	$56.51_{1.28}$	$64.20_{0.82}$
	15	$58.87_{1.20}$	$63.31_{0.91}$	$64.23_{1.70}$	$61.31_{1.08}$	$56.26_{1.15}$	$56.83_{1.41}$	$63.71_{0.45}$
	25	$59.33_{1.87}$	$63.84_{0.55}$	$64.18_{1.35}$	$61.09_{0.78}$	$56.72_{1.23}$	$56.42_{1.78}$	$64.03_{0.72}$
150	5	$59.23_{1.05}$	$64.44_{1.94}$	$64.09_{2.31}$	$62.62_{1.77}$	$55.24_{1.20}$	$56.44_{1.74}$	$63.36_{0.72}$
	10	$58.78_{1.13}$	$64.44_{1.51}$	$64.13_{2.38}$	$61.95_{1.47}$	$56.22_{1.08}$	$57.25_{1.82}$	$63.55_{0.96}$
	15	$59.28_{1.70}$	$64.25_{3.08}$	$63.72_{2.30}$	$62.03_{0.96}$	$55.85_{1.40}$	$56.57_{1.38}$	$63.53_{1.09}$
	25	$58.34_{1.39}$	$63.74_{2.45}$	$64.80_{2.16}$	$61.92_{1.05}$	$55.74_{1.03}$	$57.97_{2.15}$	$63.79_{0.71}$
200	5	$58.53_{0.59}$	$64.19_{2.65}$	$63.73_{1.11}$	$62.17_{0.77}$	$55.12_{1.17}$	$57.21_{1.98}$	$63.44_{0.72}$
	10	$58.03_{1.20}$	$63.94_{1.47}$	$63.90_{1.40}$	$62.26_{2.08}$	$55.35_{1.52}$	$56.66_{1.40}$	$63.43_{0.96}$
	15	$58.67_{1.37}$	$63.61_{1.96}$	$63.88_{1.48}$	$61.81_{1.87}$	$55.35_{1.16}$	$57.10_{1.94}$	$63.64_{0.60}$
	25	$58.71_{1.48}$	$63.80_{1.02}$	$63.74_{2.10}$	$62.10_{2.03}$	$56.05_{1.38}$	$57.29_{1.05}$	$63.15_{0.64}$

Table 6: **Semi-supervised (GraphSAGE)** Node Classification in the **Multilingual EN Corpus - German**: Macro F1 scores for the target labels *Hostility, Initiating Problem, Intolerance, Threat, Polarization, Solution,* and *Topic*. Results are reported using *In-group* and *Out-group* node attributes, varying the **number of node features** considered per node (Max EN) and the **number of neighbors** (k) used to construct the graph.

Max	K	Hostility	Initiating Problem	Intolerance	Threat	Polarization	Solution	Topic
0	5	$56.26_{1.84}$	$63.46_{1.40}$	$63.16_{1.31}$	$54.28_{0.44}$	$55.41_{2.09}$	$55.43_{0.54}$	$67.43_{1.35}$
	10	$55.60_{2.14}$	$63.28_{2.33}$	$62.97_{1.22}$	$54.39_{0.53}$	$55.34_{1.78}$	$55.98_{0.79}$	$67.90_{1.35}$
	15	$55.46_{1.73}$	$62.05_{2.49}$	$62.77_{0.79}$	$54.49_{0.45}$	$56.38_{1.08}$	$55.69_{0.55}$	$69.33_{2.19}$
	25	$54.93_{1.94}$	$63.27_{2.00}$	$62.61_{0.98}$	$54.39_{0.53}$	$55.46_{2.11}$	$56.02_{1.05}$	$68.97_{1.38}$
10	5	$56.69_{3.52}$	$57.67_{1.17}$	$62.87_{2.28}$	$53.81_{2.45}$	$52.82_{0.82}$	$54.29_{3.04}$	$66.75_{2.85}$
	10	$56.57_{2.42}$	$57.96_{0.93}$	$61.90_{1.26}$	$53.64_{2.86}$	$54.13_{0.62}$	$55.52_{1.87}$	$69.11_{2.16}$
	15	$58.22_{1.97}$	$57.69_{0.64}$	$62.38_{1.14}$	$57.42_{1.91}$	$54.03_{0.52}$	$56.28_{1.54}$	$68.32_{3.13}$
	25	$58.02_{2.95}$	$58.35_{0.63}$	$61.85_{1.84}$	$58.00_{1.55}$	$52.75_{1.17}$	$55.72_{0.89}$	$67.47_{2.64}$
50	5	$56.13_{3.04}$	$61.59_{2.26}$	$61.26_{2.17}$	$57.05_{3.45}$	$55.52_{0.95}$	$56.00_{2.17}$	$68.54_{1.30}$
	10	$55.71_{2.06}$	$61.00_{1.42}$	$62.45_{1.04}$	$55.39_{1.92}$	$54.83_{1.62}$	$56.60_{2.05}$	$69.83_{1.78}$
	15	$55.82_{2.01}$	$61.17_{1.93}$	$61.37_{1.74}$	$56.59_{3.37}$	$54.04_{1.40}$	$56.97_{2.84}$	$71.05_{1.43}$
	25	$55.12_{2.60}$	$60.91_{2.34}$	$60.60_{1.82}$	$56.08_{2.08}$	$53.77_{0.89}$	$55.71_{0.77}$	$69.75_{1.55}$
100	5	$57.45_{2.54}$	$61.51_{2.07}$	$61.91_{2.02}$	$56.15_{1.91}$	$55.28_{1.27}$	$57.55_{1.84}$	$67.92_{2.31}$
	10	$56.29_{3.24}$	$60.26_{2.34}$	$61.75_{2.23}$	$55.43_{1.22}$	$54.62_{0.81}$	$57.03_{1.87}$	$69.09_{2.06}$
	15	$55.33_{2.09}$	$61.04_{1.92}$	$62.87_{1.77}$	$55.34_{1.22}$	$55.24_{1.40}$	$56.50_{1.36}$	$68.91_{0.80}$
	25	$54.95_{2.07}$	$61.06_{1.48}$	$62.68_{1.67}$	$55.97_{2.40}$	$54.40_{1.28}$	$56.63_{1.27}$	$69.01_{0.88}$
150	5	$56.50_{1.36}$	$63.03_{1.23}$	$62.03_{2.61}$	$55.10_{2.43}$	$54.84_{0.92}$	$55.83_{0.59}$	$68.30_{2.41}$
	10	$56.18_{1.10}$	$62.08_{2.70}$	$62.45_{1.11}$	$53.92_{1.35}$	$54.53_{1.05}$	$56.18_{1.20}$	$67.58_{2.02}$
	15	$56.05_{1.17}$	$61.62_{1.83}$	$62.87_{1.79}$	$54.83_{1.46}$	$54.56_{1.21}$	$56.06_{1.33}$	$69.37_{2.17}$
	25	$55.44_{1.08}$	$61.53_{2.35}$	$61.85_{1.84}$	$55.01_{1.34}$	$54.90_{2.34}$	$56.53_{1.56}$	$67.67_{1.54}$
200	5	$56.53_{1.58}$	$62.65_{1.39}$	$62.12_{1.77}$	$54.76_{1.76}$	$56.04_{1.81}$	$55.46_{0.82}$	$68.23_{1.82}$
	10	$55.93_{1.75}$	$62.47_{1.51}$	$62.61_{0.98}$	$54.29_{0.67}$	$55.33_{1.57}$	$56.26_{0.73}$	$68.62_{1.81}$
	15	$55.71_{1.61}$	$61.98_{1.97}$	$62.82_{1.38}$	$54.13_{0.98}$	$55.17_{1.06}$	$55.81_{0.38}$	$68.85_{2.12}$
	25	$55.60_{1.50}$	$62.18_{1.83}$	$63.14_{1.15}$	$54.18_{1.01}$	$55.48_{1.27}$	$56.65_{1.42}$	$69.17_{2.38}$

Table 7: **Semi-supervised (GraphSAGE)** Node Classification in the **Multilingual EN Corpus - French**: Macro F1 scores for the target labels *Hostility, Initiating Problem, Intolerance, Threat, Polarization, Solution,* and *Topic*. Results are reported using *In-group* and *Out-group* node attributes, varying the **number of node features** considered per node (Max EN) and the **number of neighbors** (k) used to construct the graph.

Max	K	Hostility	Initiating Problem	Intolerance	Threat	Polarization	Solution	Topic
0	5	$57.91_{1.50}$	$58.21_{1.76}$	$62.74_{1.46}$	$59.41_{1.08}$	$65.29_{1.30}$	$73.12_{3.58}$	$57.97_{2.12}$
	10	$58.14_{0.94}$	$59.18_{0.44}$	$62.95_{0.74}$	$59.97_{0.69}$	$65.44_{1.10}$	$73.49_{5.93}$	$56.61_{2.54}$
	15	$58.11_{1.04}$	$59.05_{1.57}$	$63.91_{1.61}$	$59.31_{0.83}$	$64.78_{0.95}$	$75.06_{5.54}$	$58.56_{1.66}$
	25	$57.52_{0.70}$	$59.50_{1.28}$	$64.17_{0.83}$	$59.71_{2.28}$	$65.18_{0.31}$	$73.42_{5.81}$	$57.33_{1.08}$
10	5	$59.44_{1.07}$	$57.02_{1.36}$	$60.56_{0.44}$	$61.87_{0.85}$	$65.06_{1.12}$	$74.65_{5.33}$	$60.94_{1.57}$
	10	$59.01_{0.42}$	$58.00_{1.17}$	$62.02_{2.01}$	$60.45_{1.54}$	$64.78_{1.18}$	$74.50_{4.06}$	$59.85_{1.60}$
	15	$58.54_{1.76}$	$57.49_{0.97}$	$63.46_{1.34}$	$60.53_{0.67}$	$64.14_{1.71}$	$74.72_{2.63}$	$59.69_{0.84}$
	25	$58.71_{0.89}$	$57.09_{0.51}$	$62.98_{2.38}$	$58.86_{1.06}$	$64.02_{1.61}$	$75.43_{4.04}$	$59.01_{1.35}$
50	5	$62.18_{0.78}$	$57.61_{1.97}$	$64.55_{0.99}$	$61.12_{0.42}$	$67.36_{1.26}$	$74.89_{4.71}$	$61.47_{1.90}$
	10	$61.31_{2.12}$	$57.26_{1.33}$	$63.73_{0.31}$	$60.13_{0.71}$	$66.48_{0.89}$	$76.27_{5.23}$	$60.46_{1.51}$
	15	$60.92_{0.88}$	$56.40_{1.52}$	$64.44_{0.91}$	$60.86_{1.22}$	$67.27_{1.22}$	$75.45_{3.86}$	$61.67_{1.01}$
	25	$61.03_{1.22}$	$55.73_{1.14}$	$63.17_{0.52}$	$60.25_{0.90}$	$66.34_{0.77}$	$73.37_{6.97}$	$60.15_{1.36}$
100	5	$59.74_{1.27}$	$58.05_{1.36}$	$63.23_{2.12}$	$60.18_{1.33}$	$66.94_{1.48}$	$78.78_{7.24}$	$60.23_{1.48}$
	10	$60.04_{0.91}$	$58.26_{1.69}$	$63.76_{1.54}$	$61.14_{1.15}$	$66.33_{1.14}$	$76.83_{9.42}$	$59.69_{0.49}$
	15	$59.69_{1.41}$	$57.78_{1.74}$	$62.07_{0.61}$	$60.46_{0.92}$	$67.83_{2.86}$	$77.10_{5.14}$	$59.45_{1.30}$
	25	$60.02_{1.11}$	$58.75_{1.60}$	$62.73_{1.89}$	$60.14_{2.03}$	$67.08_{1.33}$	$75.19_{2.96}$	$59.85_{1.09}$
150	5	$59.34_{1.68}$	$57.98_{1.87}$	$64.04_{1.71}$	$60.05_{1.60}$	$66.09_{1.74}$	$78.90_{5.23}$	$60.47_{1.25}$
	10	$59.74_{1.14}$	$57.69_{1.56}$	$63.09_{1.65}$	$59.92_{1.47}$	$66.92_{0.83}$	$78.48_{5.37}$	$60.01_{0.65}$
	15	$59.76_{1.21}$	$57.83_{1.68}$	$62.14_{0.89}$	$60.21_{1.48}$	$66.03_{1.35}$	$76.51_{5.36}$	$60.10_{1.30}$
	25	$59.49_{1.50}$	$57.89_{0.82}$	$62.69_{1.19}$	$60.61_{0.84}$	$66.67_{1.67}$	$76.21_{6.00}$	$59.02_{0.81}$
200	5	$59.57_{1.08}$	$58.41_{2.36}$	$62.41_{1.17}$	$59.87_{2.38}$	$64.78_{0.94}$	$78.23_{7.67}$	$59.79_{0.77}$
	10	$58.87_{0.59}$	$58.72_{2.05}$	$63.34_{0.86}$	$59.37_{0.85}$	$65.76_{1.52}$	$78.21_{6.84}$	$59.15_{0.65}$
	15	$59.21_{0.95}$	$58.34_{1.91}$	$62.96_{0.79}$	$60.14_{1.52}$	$65.27_{1.10}$	$76.40_{7.77}$	$58.99_{1.43}$
	25	$60.40_{1.82}$	$57.98_{1.41}$	$62.71_{0.85}$	$60.13_{1.63}$	$66.46_{0.69}$	$76.17_{6.60}$	$58.37_{1.11}$

Table 8: **Semi-supervised (GraphSAGE)** Node Classification in the **Multilingual EN Corpus - Cypriot Greek**: Macro F1 scores for the target labels *Hostility, Initiating Problem, Intolerance, Threat, Polarization, Solution,* and *Topic*. Results are reported using *In-group* and *Out-group* node attributes, varying the **number of node features** considered per node (Max EN) and the **number of neighbors** (k) used to construct the graph.

Max	K	Hostility	Initiating Problem	Intolerance	Threat	Polarization	Solution	Topic
0	5	64.51$_{1.20}$	62.02$_{0.96}$	58.88$_{1.08}$	56.27$_{1.04}$	63.63$_{1.51}$	54.65$_{1.59}$	60.64$_{2.04}$
	10	64.92$_{1.35}$	62.14$_{0.55}$	58.52$_{1.03}$	57.75$_{1.68}$	64.67$_{1.51}$	54.53$_{1.51}$	61.89$_{1.20}$
	15	65.11$_{1.59}$	62.00$_{0.94}$	58.53$_{0.77}$	57.54$_{0.92}$	63.29$_{1.38}$	53.91$_{1.97}$	62.19$_{2.90}$
	25	65.01$_{1.59}$	61.50$_{0.71}$	58.06$_{1.16}$	57.67$_{1.05}$	62.71$_{1.04}$	53.72$_{2.04}$	60.93$_{0.30}$
10	5	63.59$_{1.64}$	58.37$_{1.53}$	58.46$_{1.22}$	55.01$_{1.36}$	62.04$_{0.79}$	51.81$_{1.85}$	68.06$_{2.37}$
	10	63.17$_{1.56}$	58.72$_{0.83}$	56.14$_{2.31}$	54.55$_{1.85}$	61.73$_{1.46}$	50.90$_{0.52}$	66.56$_{1.59}$
	15	63.24$_{1.59}$	56.96$_{1.31}$	56.88$_{1.16}$	53.17$_{1.45}$	61.34$_{1.81}$	50.82$_{1.15}$	67.67$_{1.75}$
	25	64.39$_{1.67}$	55.83$_{1.03}$	55.54$_{0.67}$	52.88$_{1.42}$	60.48$_{0.53}$	50.46$_{1.80}$	66.66$_{0.91}$
50	5	63.25$_{1.07}$	60.37$_{0.58}$	58.94$_{2.71}$	55.45$_{0.69}$	62.44$_{1.23}$	53.87$_{2.40}$	66.01$_{1.57}$
	10	64.10$_{1.04}$	59.85$_{0.46}$	59.09$_{1.07}$	55.06$_{0.91}$	61.58$_{1.28}$	53.84$_{1.70}$	66.27$_{2.05}$
	15	64.80$_{1.27}$	59.95$_{0.71}$	58.72$_{1.13}$	55.46$_{1.52}$	61.85$_{1.99}$	52.93$_{0.92}$	66.76$_{2.86}$
	25	64.78$_{0.95}$	60.26$_{1.28}$	57.89$_{2.65}$	55.32$_{1.14}$	62.58$_{2.05}$	53.32$_{1.42}$	65.89$_{1.79}$
100	5	64.27$_{1.39}$	62.28$_{1.15}$	58.93$_{1.97}$	56.37$_{1.67}$	63.16$_{2.47}$	53.20$_{2.40}$	64.08$_{1.47}$
	10	63.82$_{0.45}$	61.76$_{0.79}$	58.52$_{2.49}$	56.70$_{1.10}$	62.65$_{2.05}$	54.02$_{2.74}$	63.77$_{1.32}$
	15	64.50$_{0.83}$	61.75$_{1.36}$	58.32$_{2.29}$	56.01$_{1.68}$	62.64$_{1.54}$	52.33$_{1.31}$	62.98$_{3.24}$
	25	65.05$_{0.92}$	61.01$_{0.93}$	58.26$_{1.33}$	57.16$_{0.60}$	62.67$_{1.62}$	53.36$_{2.02}$	64.90$_{1.32}$
150	5	63.75$_{1.32}$	61.74$_{1.39}$	58.51$_{2.08}$	56.71$_{1.06}$	62.84$_{1.70}$	53.58$_{2.75}$	62.29$_{1.17}$
	10	64.17$_{1.29}$	62.24$_{1.53}$	58.63$_{1.69}$	56.28$_{0.99}$	61.93$_{2.10}$	53.71$_{2.22}$	62.46$_{0.57}$
	15	64.99$_{1.35}$	62.57$_{2.23}$	58.49$_{1.85}$	57.71$_{1.51}$	62.26$_{2.42}$	53.80$_{2.57}$	63.75$_{3.73}$
	25	64.87$_{0.90}$	62.70$_{1.56}$	57.06$_{0.92}$	56.62$_{1.36}$	62.39$_{1.56}$	53.20$_{2.43}$	61.50$_{1.45}$
200	5	64.82$_{2.03}$	62.04$_{1.61}$	59.33$_{1.45}$	56.16$_{1.61}$	62.19$_{1.50}$	53.83$_{2.49}$	60.81$_{2.70}$
	10	64.39$_{1.32}$	61.62$_{1.41}$	59.51$_{1.22}$	56.85$_{1.65}$	62.59$_{1.30}$	52.92$_{2.32}$	61.53$_{1.49}$
	15	65.02$_{1.65}$	61.63$_{0.55}$	58.70$_{1.54}$	56.92$_{2.39}$	63.32$_{2.72}$	53.15$_{2.13}$	61.08$_{1.17}$
	25	65.35$_{1.49}$	62.46$_{0.68}$	58.39$_{0.75}$	56.65$_{1.94}$	63.19$_{1.81}$	52.95$_{2.35}$	61.09$_{2.20}$

Table 9: **Semi-supervised (GraphSAGE)** Node Classification in the **Multilingual EN Corpus** - **Slovenian**: Macro F1 scores for the target labels *Hostility, Initiating Problem, Intolerance, Threat, Polarization, Solution,* and *Topic*. Results are reported using *In-group* and *Out-group* node attributes, varying the **number of node features** considered per node (Max EN) and the **number of neighbors** (k) used to construct the graph.

Max	K	Hostility	Initiating Problem	Intolerance	Threat	Polarization	Solution	Topic
0	5	59.55$_{1.55}$	63.36$_{2.30}$	57.44$_{4.06}$	57.90$_{2.49}$	53.53$_{1.82}$	57.14$_{1.63}$	68.23$_{1.60}$
	10	60.24$_{2.78}$	62.82$_{2.18}$	55.96$_{2.70}$	59.45$_{1.75}$	50.53$_{0.82}$	58.07$_{2.37}$	70.01$_{1.36}$
	15	63.59$_{1.32}$	**65.22**$_{2.84}$	61.44$_{3.72}$	58.10$_{2.13}$	54.36$_{2.21}$	56.79$_{2.54}$	69.67$_{1.17}$
	25	62.79$_{3.29}$	62.78$_{3.07}$	60.30$_{2.94}$	59.88$_{3.17}$	52.84$_{4.13}$	61.43$_{2.00}$	**70.18**$_{0.82}$
10	5	59.15$_{1.88}$	54.32$_{2.03}$	59.60$_{2.69}$	59.31$_{3.53}$	52.04$_{1.87}$	56.45$_{1.93}$	65.00$_{2.11}$
	10	61.27$_{2.49}$	61.80$_{1.40}$	57.35$_{1.80}$	60.76$_{1.22}$	54.92$_{3.66}$	58.09$_{2.44}$	67.01$_{3.12}$
	15	**63.71**$_{2.55}$	60.29$_{1.51}$	56.74$_{1.90}$	60.32$_{3.03}$	56.37$_{3.50}$	57.34$_{2.01}$	65.81$_{1.70}$
	25	61.99$_{1.88}$	61.70$_{1.80}$	58.29$_{2.95}$	63.12$_{1.86}$	55.30$_{2.01}$	54.68$_{2.03}$	66.98$_{1.80}$
50	5	62.20$_{1.65}$	61.95$_{2.58}$	59.63$_{3.42}$	58.16$_{3.34}$	55.86$_{2.17}$	57.46$_{2.56}$	66.04$_{2.22}$
	10	62.39$_{0.87}$	61.34$_{1.47}$	60.39$_{4.91}$	62.37$_{2.12}$	57.38$_{1.57}$	54.90$_{2.15}$	65.28$_{2.60}$
	15	60.83$_{2.50}$	63.52$_{3.02}$	60.23$_{1.23}$	60.00$_{1.92}$	54.48$_{4.40}$	57.73$_{1.50}$	67.28$_{0.82}$
	25	61.16$_{1.52}$	62.41$_{2.92}$	58.28$_{2.15}$	**63.03**$_{2.35}$	57.68$_{2.63}$	58.23$_{2.57}$	66.45$_{1.71}$
100	5	58.61$_{2.40}$	56.39$_{2.18}$	58.01$_{3.25}$	59.96$_{1.15}$	55.19$_{2.37}$	56.43$_{1.74}$	66.42$_{2.56}$
	10	62.76$_{1.85}$	59.44$_{2.16}$	59.20$_{3.07}$	61.48$_{3.25}$	56.46$_{3.16}$	59.47$_{2.56}$	66.86$_{1.90}$
	15	60.78$_{2.99}$	59.73$_{2.53}$	61.50$_{2.21}$	62.55$_{1.50}$	53.86$_{1.88}$	57.92$_{2.28}$	66.62$_{1.90}$
	25	63.54$_{1.82}$	60.45$_{1.40}$	**61.89**$_{1.81}$	62.20$_{3.19}$	55.06$_{1.69}$	**61.56** $_{3.57}$	67.17$_{0.99}$
150	5	58.61$_{2.82}$	55.01$_{0.76}$	61.20$_{2.39}$	59.49$_{1.83}$	56.54$_{1.66}$	56.21$_{2.04}$	67.07$_{2.11}$
	10	60.99$_{2.02}$	58.79$_{2.01}$	58.58$_{4.25}$	61.94$_{1.01}$	55.45$_{2.13}$	58.70$_{2.51}$	67.48$_{3.66}$
	15	62.51$_{2.57}$	62.04$_{2.86}$	57.45$_{2.06}$	62.95$_{1.11}$	55.67$_{1.93}$	57.82$_{1.41}$	65.04$_{2.36}$
	25	62.92$_{2.62}$	63.37$_{1.85}$	57.66$_{2.11}$	63.47$_{3.28}$	54.61$_{3.18}$	59.60$_{0.69}$	67.45$_{3.01}$
200	5	60.93$_{2.47}$	61.45$_{1.35}$	59.54$_{2.66}$	60.07$_{0.72}$	58.38$_{2.03}$	59.58$_{1.73}$	63.64$_{1.91}$
	10	61.23$_{2.46}$	62.32$_{1.85}$	58.53$_{3.49}$	61.06$_{2.85}$	52.61$_{1.80}$	52.98$_{2.27}$	66.41$_{1.25}$
	15	61.85$_{2.34}$	59.30$_{2.54}$	59.70$_{3.83}$	62.46$_{1.86}$	**56.83**$_{4.04}$	59.29$_{1.40}$	67.96$_{2.42}$
	25	62.68$_{2.28}$	60.38$_{2.07}$	55.62$_{4.09}$	61.55$_{2.02}$	55.90$_{2.22}$	58.20$_{2.74}$	64.60$_{1.87}$

Table 10: **Contrastive Learning** for Node Classification in the **Multilingual EN Corpus** - **German**: Macro F1 scores for the target labels *Hostility, Initiating Problem, Intolerance, Threat, Polarization, Solution,* and *Topic*. Results are reported using *In-group* and *Out-group* node attributes, varying the **number of node features** considered per node (Max EN) and the **number of neighbors** (k) used to construct the graph.

Max	K	Hostility	Initiating Problem	Intolerance	Threat	Polarization	Solution	Topic
0	5	$54.83_{3.04}$	$60.19_{1.93}$	$49.45_{3.08}$	$47.14_{0.26}$	$46.67_{2.10}$	$49.14_{2.49}$	$56.47_{5.00}$
	10	$53.27_{4.11}$	$62.36_{3.24}$	$48.07_{1.97}$	$47.21_{0.48}$	$49.38_{0.97}$	$51.29_{3.81}$	$50.92_{1.57}$
	15	$56.22_{5.59}$	$60.94_{3.24}$	$49.41_{3.96}$	$47.45_{0.23}$	$47.04_{1.59}$	$49.60_{2.48}$	$51.70_{1.64}$
	25	$51.14_{3.23}$	$58.21_{1.66}$	$52.52_{1.51}$	$47.80_{0.28}$	$49.17_{0.67}$	$48.21_{0.38}$	$51.10_{4.83}$
10	5	$56.62_{7.83}$	$60.24_{2.45}$	$50.20_{2.78}$	$\mathbf{52.99}_{1.68}$	$\mathbf{51.43}_{2.57}$	$51.91_{5.59}$	$57.89_{7.07}$
	10	$55.10_{6.42}$	$61.13_{3.13}$	$49.60_{2.72}$	$51.39_{4.42}$	$47.75_{1.66}$	$\mathbf{57.05}_{2.01}$	$51.22_{1.21}$
	15	$51.38_{3.62}$	$59.67_{4.42}$	$51.75_{4.23}$	$49.70_{2.74}$	$48.03_{1.97}$	$51.97_{3.17}$	$54.70_{3.56}$
	25	$50.60_{3.79}$	$58.61_{2.05}$	$52.20_{5.46}$	$49.44_{3.23}$	$45.84_{0.65}$	$54.42_{3.54}$	$51.45_{2.38}$
50	5	$48.66_{2.66}$	$57.52_{2.34}$	$53.02_{2.23}$	$51.55_{4.60}$	$50.26_{2.63}$	$51.26_{3.86}$	$62.16_{4.37}$
	10	$55.80_{5.77}$	$56.43_{1.81}$	$55.20_{4.61}$	$49.57_{3.16}$	$47.01_{1.63}$	$49.96_{3.29}$	$61.80_{4.36}$
	15	$54.86_{4.41}$	$59.43_{1.17}$	$49.06_{3.28}$	$52.16_{3.74}$	$48.73_{1.94}$	$55.82_{5.51}$	$57.71_{6.36}$
	25	$51.03_{4.05}$	$58.06_{4.11}$	$56.44_{3.62}$	$49.33_{3.29}$	$49.21_{2.31}$	$51.57_{3.55}$	$55.80_{6.91}$
100	5	$\mathbf{57.02}_{1.66}$	$53.41_{1.38}$	$52.00_{1.95}$	$48.11_{0.17}$	$49.97_{4.18}$	$54.43_{3.83}$	$58.22_{6.26}$
	10	$52.27_{2.12}$	$\mathbf{62.69}_{3.12}$	$50.38_{2.94}$	$52.07_{3.46}$	$47.71_{2.18}$	$49.45_{1.98}$	$53.57_{2.53}$
	15	$48.66_{2.36}$	$56.78_{2.68}$	$52.89_{1.27}$	$51.29_{3.86}$	$48.01_{2.13}$	$50.04_{3.33}$	$52.42_{0.89}$
	25	$51.56_{3.49}$	$60.50_{1.69}$	$56.10_{1.97}$	$47.65_{0.74}$	$49.88_{0.54}$	$52.70_{3.98}$	$52.85_{0.59}$
150	5	$56.50_{5.72}$	$56.46_{1.25}$	$55.21_{6.80}$	$47.63_{0.26}$	$49.82_{2.59}$	$50.12_{3.76}$	$51.64_{2.52}$
	10	$53.43_{3.82}$	$56.02_{1.95}$	$50.49_{4.38}$	$51.31_{3.07}$	$50.96_{1.19}$	$52.00_{2.95}$	$53.50_{6.90}$
	15	$47.66_{0.33}$	$60.07_{1.63}$	$\mathbf{57.48}_{4.93}$	$52.73_{4.05}$	$48.07_{3.50}$	$48.86_{2.33}$	$\mathbf{65.01}_{3.19}$
	25	$53.19_{4.82}$	$59.71_{2.76}$	$55.28_{5.00}$	$51.05_{3.19}$	$48.81_{1.03}$	$49.53_{2.82}$	$56.81_{5.75}$
200	5	$50.33_{2.96}$	$57.43_{1.84}$	$51.68_{4.37}$	$48.02_{0.93}$	$48.59_{1.43}$	$54.28_{3.03}$	$56.11_{2.36}$
	10	$56.13_{2.62}$	$56.75_{2.40}$	$55.04_{2.71}$	$48.73_{2.10}$	$46.66_{1.85}$	$54.28_{5.05}$	$54.54_{7.10}$
	15	$52.55_{2.71}$	$59.19_{2.81}$	$51.68_{2.61}$	$49.05_{2.20}$	$46.59_{1.66}$	$52.72_{3.89}$	$58.27_{7.60}$
	25	$51.99_{2.40}$	$60.72_{1.12}$	$51.31_{3.61}$	$47.52_{0.19}$	$48.02_{2.27}$	$51.41_{4.00}$	$52.18_{1.68}$

Table 11: **Contrastive Learning** Node Classification in the **Multilingual EN Corpus** - **French**: Macro F1 scores for the target labels *Hostility, Initiating Problem, Intolerance, Threat, Polarization, Solution,* and *Topic*. Results are reported using *In-group* and *Out-group* node attributes, varying the **number of node features** considered per node (Max EN) and the **number of neighbors** (k) used to construct the graph.

Max	K	Hostility	Initiating Problem	Intolerance	Threat	Polarization	Solution	Topic
0	5	$56.12_{1.03}$	$72.18_{0.74}$	$56.61_{4.42}$	$65.52_{4.26}$	$64.82_{2.01}$	$62.29_{2.07}$	$74.91_{0.68}$
	10	$56.21_{2.15}$	$70.94_{1.49}$	$58.11_{2.73}$	$65.35_{2.05}$	$65.01_{2.54}$	$69.06_{7.67}$	$75.29_{3.24}$
	15	$\mathbf{58.50}_{3.12}$	$72.00_{2.46}$	$57.93_{1.23}$	$68.05_{1.56}$	$65.46_{2.47}$	$74.49_{4.40}$	$75.86_{0.69}$
	25	$57.33_{2.86}$	$71.70_{1.49}$	$62.25_{4.31}$	$67.16_{3.64}$	$63.82_{2.98}$	$63.38_{5.15}$	$75.60_{1.41}$
10	5	$54.99_{1.41}$	$71.87_{2.72}$	$61.77_{1.60}$	$63.83_{3.13}$	$64.13_{1.91}$	$63.58_{1.81}$	$73.08_{2.13}$
	10	$54.73_{0.80}$	$71.65_{1.12}$	$63.48_{2.14}$	$65.75_{3.36}$	$66.48_{2.98}$	$62.86_{2.77}$	$75.65_{0.95}$
	15	$56.16_{2.97}$	$69.50_{2.24}$	$60.68_{3.53}$	$63.29_{2.59}$	$\mathbf{67.14}_{0.72}$	$71.80_{5.38}$	$75.26_{2.52}$
	25	$58.29_{1.72}$	$70.33_{3.44}$	$60.63_{3.06}$	$67.44_{1.67}$	$64.83_{1.06}$	$70.36_{3.65}$	$\mathbf{76.15}_{2.10}$
50	5	$52.11_{2.03}$	$69.31_{2.01}$	$63.12_{3.04}$	$62.95_{2.99}$	$60.88_{1.69}$	$69.29_{3.80}$	$75.41_{0.62}$
	10	$52.75_{1.81}$	$68.75_{1.50}$	$59.45_{0.93}$	$\mathbf{68.64}_{2.39}$	$60.94_{2.08}$	$76.59_{3.28}$	$75.53_{1.27}$
	15	$54.69_{0.71}$	$70.83_{1.28}$	$56.75_{2.72}$	$66.19_{0.58}$	$62.75_{1.95}$	$74.65_{2.52}$	$74.89_{1.92}$
	25	$57.65_{4.28}$	$69.71_{1.25}$	$60.27_{1.27}$	$65.00_{2.90}$	$65.35_{2.49}$	$73.83_{2.15}$	$73.93_{1.57}$
100	5	$56.25_{2.62}$	$70.02_{2.94}$	$\mathbf{64.55}_{2.13}$	$63.58_{1.13}$	$59.10_{3.28}$	$67.05_{3.86}$	$73.21_{2.83}$
	10	$53.44_{3.01}$	$70.73_{3.00}$	$62.47_{1.92}$	$66.63_{2.28}$	$61.52_{1.99}$	$73.70_{2.33}$	$75.40_{2.15}$
	15	$54.25_{2.32}$	$70.16_{2.59}$	$58.48_{3.55}$	$66.96_{2.30}$	$63.98_{2.59}$	$76.34_{6.34}$	$74.11_{1.96}$
	25	$51.72_{2.08}$	$72.31_{1.37}$	$58.77_{3.40}$	$65.25_{1.38}$	$64.72_{2.47}$	$\mathbf{76.77}_{1.58}$	$74.47_{1.66}$
150	5	$54.87_{2.07}$	$73.17_{2.41}$	$55.17_{1.58}$	$62.73_{2.20}$	$64.05_{1.51}$	$66.68_{4.64}$	$72.38_{1.79}$
	10	$60.48_{3.03}$	$71.47_{2.44}$	$58.92_{3.54}$	$65.19_{1.19}$	$66.88_{2.03}$	$71.57_{3.82}$	$74.82_{0.82}$
	15	$57.22_{3.84}$	$71.33_{0.88}$	$60.13_{2.61}$	$65.69_{2.07}$	$65.18_{2.63}$	$68.77_{4.22}$	$75.00_{0.95}$
	25	$57.16_{3.65}$	$72.44_{2.12}$	$62.04_{2.27}$	$67.81_{1.98}$	$65.50_{1.44}$	$75.04_{7.68}$	$73.92_{1.26}$
200	5	$56.03_{1.24}$	$69.23_{2.21}$	$57.71_{3.03}$	$61.73_{1.58}$	$63.20_{3.08}$	$71.59_{3.11}$	$74.67_{1.38}$
	10	$55.49_{2.43}$	$70.16_{1.78}$	$56.11_{1.90}$	$63.97_{1.67}$	$63.30_{2.87}$	$76.50_{3.38}$	$74.60_{2.14}$
	15	$57.06_{1.69}$	$\mathbf{72.82}_{1.14}$	$55.49_{2.47}$	$65.95_{1.47}$	$63.75_{2.25}$	$69.51_{7.61}$	$71.78_{1.56}$
	25	$56.41_{1.38}$	$71.12_{2.96}$	$58.70_{2.39}$	$64.92_{2.15}$	$66.57_{2.78}$	$71.86_{2.91}$	$73.86_{2.39}$

Table 12: **Contrastive Learning** Node Classification in the **Multilingual EN Corpus** - **Cypriot Greek**: Macro F1 scores for the target labels *Hostility, Initiating Problem, Intolerance, Threat, Polarization, Solution,* and *Topic*. Results are reported using *In-group* and *Out-group* node attributes, varying the **number of node features** considered per node (Max EN) and the **number of neighbors** (k) used to construct the graph.

Max	K	Hostility	Initiating Problem	Intolerance	Threat	Polarization	Solution	Topic
0	5	$57.37_{2.28}$	$60.38_{2.53}$	$59.87_{3.83}$	$58.69_{4.45}$	$60.63_{3.09}$	$49.78_{3.75}$	$61.10_{2.10}$
	10	$56.78_{2.47}$	$58.54_{1.27}$	$58.99_{2.13}$	$53.04_{3.79}$	$60.44_{2.85}$	$49.48_{2.42}$	$61.23_{1.81}$
	15	$58.40_{0.90}$	$54.59_{3.44}$	$60.83_{1.93}$	$54.65_{2.59}$	$66.22_{1.23}$	$53.35_{0.97}$	$\mathbf{65.69}_{2.74}$
	25	$\mathbf{60.29}_{1.26}$	$57.09_{2.97}$	$\mathbf{63.27}_{1.73}$	$54.81_{4.15}$	$61.30_{2.12}$	$54.65_{2.91}$	$63.62_{1.79}$
10	5	$57.35_{2.14}$	$56.32_{1.79}$	$56.74_{1.71}$	$57.56_{1.48}$	$61.43_{1.75}$	$47.73_{1.01}$	$61.57_{3.14}$
	10	$54.18_{1.82}$	$59.64_{1.19}$	$57.65_{1.89}$	$56.72_{2.16}$	$59.98_{3.14}$	$53.28_{2.51}$	$58.08_{1.30}$
	15	$51.78_{1.35}$	$59.03_{3.09}$	$59.23_{2.13}$	$59.51_{1.67}$	$60.59_{3.56}$	$54.44_{1.55}$	$62.43_{2.29}$
	25	$53.55_{0.93}$	$58.75_{1.20}$	$58.98_{2.26}$	$57.54_{1.00}$	$61.46_{1.74}$	$54.26_{3.04}$	$61.20_{2.48}$
50	5	$58.82_{2.02}$	$58.80_{1.03}$	$56.35_{2.61}$	$56.95_{3.35}$	$61.07_{2.69}$	$52.01_{1.81}$	$59.35_{2.06}$
	10	$55.85_{2.62}$	$57.08_{2.67}$	$57.37_{2.69}$	$58.82_{2.33}$	$61.63_{3.23}$	$53.34_{4.42}$	$60.91_{2.69}$
	15	$54.27_{2.25}$	$59.63_{2.62}$	$60.06_{2.43}$	$57.67_{3.43}$	$58.35_{1.83}$	$53.59_{1.88}$	$62.55_{2.85}$
	25	$54.09_{1.46}$	$57.01_{3.49}$	$60.50_{1.22}$	$56.33_{2.91}$	$58.37_{2.58}$	$\mathbf{54.86}_{3.05}$	$63.44_{3.57}$
100	5	$59.54_{2.42}$	$\mathbf{60.77}_{2.01}$	$60.29_{0.78}$	$57.99_{2.25}$	$61.28_{2.71}$	$52.61_{2.79}$	$60.70_{1.29}$
	10	$54.03_{2.32}$	$58.42_{2.01}$	$58.08_{1.48}$	$58.78_{3.26}$	$61.69_{3.39}$	$53.95_{1.98}$	$63.41_{2.62}$
	15	$54.80_{2.97}$	$56.88_{2.16}$	$59.36_{2.66}$	$\mathbf{60.53}_{4.08}$	$\mathbf{64.08}_{1.98}$	$55.84_{2.09}$	$63.43_{1.37}$
	25	$56.29_{0.88}$	$60.15_{2.18}$	$58.67_{1.87}$	$56.97_{2.49}$	$63.27_{1.39}$	$54.52_{2.05}$	$61.59_{1.77}$
150	5	$56.18_{2.57}$	$58.59_{1.74}$	$56.88_{2.24}$	$57.77_{1.80}$	$62.97_{1.79}$	$51.89_{1.08}$	$62.23_{0.88}$
	10	$53.87_{1.29}$	$57.68_{1.97}$	$57.64_{2.47}$	$53.13_{1.74}$	$60.71_{2.48}$	$52.03_{3.24}$	$65.46_{2.52}$
	15	$52.34_{3.08}$	$59.61_{0.82}$	$59.45_{1.18}$	$54.93_{2.37}$	$57.99_{3.48}$	$53.74_{2.24}$	$65.08_{2.50}$
	25	$54.49_{1.28}$	$56.88_{1.45}$	$57.83_{1.99}$	$57.45_{3.08}$	$60.33_{2.18}$	$49.56_{1.44}$	$64.21_{3.08}$
200	5	$53.80_{1.44}$	$59.52_{2.49}$	$57.43_{0.94}$	$54.79_{1.83}$	$59.84_{1.73}$	$52.77_{1.96}$	$59.47_{2.08}$
	10	$55.59_{2.47}$	$57.71_{1.26}$	$57.68_{2.55}$	$56.07_{2.98}$	$60.52_{1.76}$	$54.42_{2.05}$	$59.52_{3.06}$
	15	$58.50_{1.75}$	$59.58_{1.77}$	$56.78_{1.53}$	$56.70_{3.05}$	$60.80_{1.48}$	$56.33_{2.18}$	$60.84_{1.94}$
	25	$58.01_{2.05}$	$59.05_{0.69}$	$59.05_{1.23}$	$54.15_{3.24}$	$60.55_{2.86}$	$54.61_{2.94}$	$61.97_{1.03}$

Table 13: **Contrastive Learning** Node Classification in the **Multilingual EN Corpus - Slovenian**: Macro F1 scores for the target labels *Hostility, Initiating Problem, Intolerance, Threat, Polarization, Solution,* and *Topic*. Results are reported using *In-group* and *Out-group* node attributes, varying the **number of node features** considered per node (Max EN) and the **number of neighbors** (k) used to construct the graph.