Prompting Politics: Controlled Generation of Italian Political Tweets via LoRA-Tuned LLMs

Beatrice Citterio **Filippo Focaccia** **Giulio Pirotta** **Martina Serandrei** **Tommaso Vezzoli**

Abstract

Large language models (LLMs) are impressively versatile, yet in political communication on social media they often miss the subtle interplay of ideology, topic, and tone that gives a tweet its persuasive power. In this work, we present a controlled tweet generation framework tailored for Italian political discourse, which allows manipulation of political stance, sentiment, and topic. We construct a high-quality dataset of 17,000 tweets from key Italian politicians, annotated through topic modeling, custom sentiment categories, and ideological alignment. We load two pretrained Italian LLMs from the Minerva family and we leverage Low-Rank Adaptation (LoRA) for efficient fine-tuning. These models allows us to condition generation on structured natural language prompts. Our experiments demonstrate that the fine-tuned models significantly outperform baselines in aligning output with the input sentiment, while preserving fluency and ideological consistency. This research establishes a new benchmark for text generation in the Italian political context.

1 Introduction

Large language models (LLMs) have demonstrated remarkable versatility across a wide range of natural language processing tasks, including summarization, translation, and dialogue generation. However, their effectiveness in capturing the complex nature of political communication remains limited. The persuasive power and social resonance of political messages are shaped by several factors, such as the ideological position, rhetorical tone, and subject matter. Current generative models often struggle to reproduce this layered structure, especially in multilingual and culturally specific contexts such as Italian politics.

In this work, we propose a controlled tweet generation framework specifically designed to model the Italian political landscape on Twitter. Our setting enables precise manipulation of three key dimensions of political discourse: political affiliation, sentiment, and topic. This setup allows researchers and practitioners to simulate tweets that not only sound authentic but also reflect clearly defined ideological and emotional positions.

To accomplish our goal, we create a high-quality dataset consisting of approximately 17,000 tweets posted by relevant Italian politicians. These tweets span multiple parties, ideological backgrounds, and topical domains. We annotated the dataset using unsupervised topic modeling and sentiment analysis via transfer learning. Finally, we assign to each tweet an ideology label based on the political affiliation of the writer. This enriched annotation enables fine-grained control over tweet generation and provides a valuable resource for downstream tasks.

We use two pretrained Italian language models from the Minerva family as the foundation of our generation system. To make training both efficient and scalable, we apply Low-Rank Adaptation (LoRA), a parameter-efficient fine-tuning method that enables the models to specialize on our political tweet dataset without the need for full retraining. This approach significantly reduces computational requirements while maintaining strong performance. Our models are designed to respond to structured natural language prompts, allowing users to specify the desired stance, sentiment, and topic in intuitive ways. Through this conditioning mechanism, we generate politically coherent tweets that reflect the requested attributes.

We evaluate our fine-tuned models across several dimensions, including alignment with input sentiment and topical relevance. Our results show that the LoRA-adapted Minerva models consistently outperform the pre-trained models, particularly in capturing emotional cues. The generated outputs maintain high linguistic quality while adapting effectively to diverse political tones and themes.

Ultimately, this research sets a new benchmark for controlled political text generation in the Italian Twitter framework. Beyond model performance, we highlight potential applications in simulating political conversations for educational tools, political science research, and even real-time interactive scenarios.

2 Methodology

In this section, we present the rationale behind creating our dataset and all the methods and approaches we experiment to reach our final objective.

2.1 Data Description and Processing

The data is collected from X (Twitter) by scraping tweets through a python library called *tweety* (Kharl, 2024), which reverse-engineers the Twitter frontend API. Thanks to this powerful tool, we were able to scrape 27,000 tweets by considering one or two Italian politicians for each main party. The following parties were considered: Fratelli di Italia (Meloni and La Russa), Forza Italia (Berlusconi and Tajani), Lega (Salvini), Movimento Cinque Stelle (Di Maio and Conte), Azione (Calenda), Italia Viva (Renzi), +Europa (Emma Bonino), Partito Democratico (Schlein and Letta), Europa Verde (Fratoianni), Noi con l'Italia (Lupi).

Each tweet has the following attributes:

- date
- unique ID
- URL, which allows to retrieve the original post
- text content
- number of likes
- number of retweets

Note that tweets are selected from the 1st of January 2021 up until the 31st of December 2022, so that they capture the entire electoral campaign of the September 2022 elections.

What emerges is that many tweets don't express any real political message or value, and they are mainly about promoting the politicians' presence at a talk show or at a live event. We resolve this issue using OpenAI API to label each tweet either as 'SIGNIFICANT' or 'BROADCAST' based on *ad-hoc* instructions.[1] At the end of this process, around 17,000 tweets remain, covering every politician, except for La Russa.

2.2 Topic Modeling

In this section we explain the topic modeling procedure. The first step consists in pre-processing the data, by removing URLs, mentions, hashtags, and punctuation. Then, we split the politicians in two main categories 'right wing' and 'left wing'.
The former comprises: Meloni, Salvini, Berlusconi, Lupi, and Tajani. The latter has tweets from: Letta, Conte, Calenda, Renzi, Bonino, Di Maio, Fratoianni, Bonelli, Schlein.

To infer the topics, we try the following approaches:

- **Gensim LDA model**, based on the Variational Bayes algorithm
- **MALLET** from the *little-mallet-wrapper* Python library (Antoniak, 2020), which implements LDA using Gibbs sampling;

We calculate **coherence** using the C_v score to assess the optimal number of topics on our corpus, as suggested by (Röder et al., 2015).
We find that for both datasets, MALLET suggests 18 as the optimal number of topics. This choice is confirmed by direct inspection of the data.
Finally, we label each tweet with the corresponding topic and we combine the topics into three macro categories for each wing:

- right wing topics:
  - National Values and Social Issues
  - Economic and Development Issues
  - Governance and International Relations
- left wing topics:
  - International Relations and Progressive Values
  - Economic and Social Policy
  - Governance and Political Positioning

More information can be found in the Appendix 5.

2.3 Sentiment Analysis

First, we try standard sentiment analysis procedures through pre-trained models. In particular, we

[1] To take a look at the original prompt given to the OpenAI API for the classification, refer to Appendix 5.

implement two models taken from Hugging Face website:

- **RoBERTa-based model** (*twitter-xlm-roberta-base*) (Barbieri et al., 2022), trained on ∼198M multilingual tweets (among which Italian);
- **UmBERTo-based model** (*feel-it-italian-sentiment*) (Bianchi et al., 2021), fine-tuned only on Italian tweets.

However, both models classify the tweets in a significantly unbalanced manner. For this reason, we further investigate the outputs by manually reviewing samples. This inspection reveals that the standard labels (positive, negative, and neutral) do not adequately represent the semantic complexity of our data.

Active learning We address this issue by performing active learning (Zhang et al., 2022) with task-specific labels. Further analysis shows that tweets can be categorized in four main clusters:

- neutral / informative: factual updates, announcements, or descriptive statements with no clear emotional tone or stance;
- positive / supportive / celebratory: positive endorsement or praise for allies, institutions, values, events;
- angry / critical: direct attack, blame, or disapproval, usually aimed at political opponents;
- propaganda / call to action: tweets that promote a political agenda or mobilize citizens to take part in the September 2022 elections;

Examples from each class can be found in Appendix 5.

Then, we start with the active learning procedure: we manually label 200 tweets and fine-tune a pre-trained BERT-based model (*bert-base-italian-uncased* (Schweter, 2020)) on these. Next, we perform inference on the rest of the dataset, keeping track of the average prediction confidence. The next batch to label consists of the 100 tweets on which the model is least confident. We label these tweets and repeat the process until we reach a satisfactory level of average confidence. In our case, we achieve average certainty of 73% with 700 labeled samples. Finally, we apply the fine-tuned model to the whole dataset, select samples predicted with confidence above 90% and add them to the training data. We train the model on this comprehensive corpus, apply it to the rest of the labels and achieve 78% of average confidence. From the predictions, we select tweets labeled with confidence above 80% as this allows us to retain 9606 labeled samples. For all the tweets below the confidence threshold, we set the tone label to *unknown*.

2.4 Generative Model

This section encloses the methodology of the most substantial task in this project: turn a general-purpose Italian LLM into a *prompt-conditioned* tweet generator. The model's main characteristic is obeying three control variables — party, topic, sentiment — embedded in natural language instructions.

The Minerva family seems to be the most suitable because it already tokenizes Italian language efficiently, while remaining small enough for a single GPU. Two checkpoints are explored: **350 M parameters** (baseline for commodity hardware) and **1 B parameters** (to assess scaling benefits).

For the sake of consistency, we structure the input into the *instruction–answer* format expected by the model:

```
<s[INST] Scrivi un tweet come se fossi un
politico...
Partito: {party} Argomento: {topic}
Accezione: {sentiment} [/INST] tweet</s>
```

Padding is left-aligned; every label position that belongs to the prompt is set to -100 so the loss is computed solely on the tweet tokens. We then split the corpus 70% train, 20% test and a light 10% eval slice used just as a sanity check during the training loop.

Why LoRA and how it works Full fine-tuning of 350 M+ parameters is wasteful for stylistic adaptation. **LoRA** is an advanced fine-tuning technique designed to reduce the number of trainable parameters in LLMs without significantly compromising performance. It consists in freezing the pre-trained model weights and injecting trainable rank decomposition matrices into each layer of the Transformer architecture. We set the following hyperparameters:

- $r = 8$: rank of the low-dimensional update; higher means more capacity, lower means lighter adapters;
- $\alpha = 32$: scaling factor; the effective learning rate of the adapter becomes α/r;

- **dropout** = 0.05: applied on the LoRA branch only, it regularizes the new parameters without touching the frozen base;

Targeted modules are the query, key, value, output and feed-forward projections. All weights are kept in **4-bit NF4 quantisation**[2]; the adapters themselves stay in FP16.

Training schedule We run three epochs using AdamW optimizer with learning rate = 2×10^{-4}, batch size = 4 and gradient accumulation = 8, giving an effective batch of 32 without exceeding memory. Gradient checkpointing and scaler further reduce the footprint. With these settings training the 350 M model completes in about two hour, while the 1 B model takes three and a half.

Picking the Best Checkpoint At the end of each epoch we pause, generate the validation tweets, and measure two quick n-gram metrics:

- **BLEU-2**, measuring short lexical matches,
- **ROUGE-L**, which rewards longer common subsequences and is more forgiving when wording changes;

Although these scores do not measure topic or sentiment directly, they track how fluently the model echoes the reference tweet and are therefore a handy proxy during training.

Inference For every prompt in the held-out test set we produce:

1. a tweet from the base model,
2. a tweet from the fine-tuned checkpoint,
3. the original tweet attached to the prompt.

The inference script streams the prompts in batches of 32 and generates up to 128 new tokens per prompt with temperature = 0.7. Then, it stores all outputs together with the controlling fields (party, topic, sentiment). The resulting file is the substrate for the evaluation section that follows.

In conclusion, Low-Rank Adaptation allows us to specialize Minerva with just ≈0.52% new parameters in Minerva-1B and ≈0.85% in Minerva-350M, while keeping the rest of the network frozen and quantized. The procedure is memory-light, repeatable on commodity hardware, and — most importantly — produces two comparable sets of generated tweets (base vs. tuned) keyed to the same prompts.

[2] i.e. a non-uniform 4-bit format which compared to the uniform int-4 preserves almost full precision accuracy, cutting memory by 8 times. We use the implementation from BitsAndBytes.

3 Results

To evaluate the performance of our generative model, we want to assess how the generated tweets match the prompted labels. To do so, we build two classifiers, one for the topic and one for the sentiment.

3.1 Topic Evaluation

The topic classifier is trained on the full politicians' dataset, where each observation has the following features:

- preprocessed tweet
- wing (0 corresponding to right, 1 corresponding to left)
- sentiment

We perform the rest of the analysis on the right and left wing datasets separately. We encode the target topics into a categorical variable and vectorize tweets using TFIDF with the following parameters:

- max_features=1000,
- min_df=2,
- max_df=0.8,
- ngram_range=(1, 2)

We train a random forest classifier and perform inference on the tweets generated by both finetuned and baseline models. We notice that all models produce comparable results in terms of accuracy, suggesting that fine-tuning does not enhance the ability to stick to the prompted topic. There are several reasons that explain this phenomenon. First, the labeling procedure is a non-trivial task, thus it might have produced some noisy labels. Furthermore, topics are highly specific and there may be some semantic overlaps between them.

3.2 Sentiment Evaluation

To perform sentiment evaluation we fine-tune *bert-base-italian-uncased* on the entire labeled dataset produced in Section 2.3, discarding *unknown* samples. Then, we perform inference on the generated data, filtered by labels which are not *generic*. We compute accuracy as the percentage of matches between the label given as input prompt and the one inferred by the classifier (which we consider to be the ground truth). Interestingly, for the 350M model, we witness an increase in accuracy from 24% (baseline) to 76% (finetuned), showing that our prompt-based strategy has succeeded. The 1B model shows a similar trend, going from 32% of accuracy (baseline) to 70% (finetuned). It's not surprising that the 350M finetuned model performs the best, as it involves updating the largest proportion of parameters. This allows the model to learn more effectively how to follow the given prompt - in this case, the sentiment label.

4 Related Work

This project builds upon several distinct branches of NLP research: political text generation, topic modeling, sentiment analysis, and fine-tuning techniques for language models. While much has been done in each individual area, their integration into a conditional tweet generator tailored for Italian political simulation remains underexplored. Our approach addresses this gap by combining ideological stance, topical relevance, and emotional tone into a unified generative framework.

Political tweet generation has primarily focused on stylistic emulation using general-purpose LLMs or fine-tuned GPT models. For instance, (Garvey et al., 2021) explore how political speech can be simulated using large transformers, but their analysis lacks explicit controllability over sentiment or political side. Similarly, recent models by (Russo et al., 2024) use retrieval-based methods for enhancing factuality in political generation, but still do not encode sentiment or ideological alignment as input parameters. Our model adds value by making sentiment, topic, and political stance explicit conditions for generation, improving control and interpretability.

5 Conclusion

In this work, we present a novel framework for controlled tweet generation in the context of Italian political discourse, integrating sentiment, topic, and ideological stance as explicit conditioning variables. Through the construction of a tailored dataset and the application of LoRA-based fine-tuning on two Italian language models from the Minerva family, we demonstrate that stylistically and semantically faithful political text generation is not only feasible but effective, even on relatively small models.

Our results show that the 350M fine-tuned model outperforms both larger and baseline counterparts in adhering to prompted sentiment labels, confirming the efficiency of the LoRA strategy. Although topic alignment remains a more challenging task, partly due to noisy annotations and overlapping semantic fields, the framework lays a solid foundation for future refinements.

This research contributes with a flexible and interpretable approach to political text simulation in Italian, with several potential applications.

References

Maria Antoniak. 2020. little-mallet-wrapper: A python wrapper for mallet topic modeling. Version 0.0.4. Available at: https://pypi.org/project/little-mallet-wrapper/.

Francesco Barbieri, Luis Espinosa Anke, and Jose Camacho-Collados. 2022. XLM-T: Multilingual language models in twitter for sentiment analysis and beyond. In *Proceedings of the 2022 International Conference on Language Resources and Evaluation (LREC)*, Marseille, France. European Language Resources Association (ELRA).

Federico Bianchi, Debora Nozza, and Dirk Hovy. 2021. FEEL-IT: Emotion and sentiment classification for the Italian language. In *Proceedings of the Eleventh Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, pages 76–83, Online. Association for Computational Linguistics.

Myles D. Garvey, Jim Samuel, and Alexander Pelaez. 2021. Would you please like my tweet?! an artificially intelligent, generative probabilistic, and econometric based system design for popularity-driven tweet content generation. *Decision Support Systems*, 144:113497.

Tayyab Kharl. 2024. Tweety: Reverse engineered twitter frontend api. Version 2.1. Available at: https://mahrtayyab.github.io/tweety$_{docs}/$.

Michael Röder, Andreas Both, and Alexander Hinneburg. 2015. Exploring the space of topic coherence measures. In *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining (WSDM)*, pages 399–408, Shanghai, China. ACM.

Daniel Russo, Thiemo Wambsganss, and Alexander Maedche. 2024. Face the facts! evaluating rag-based fact-checking pipelines in realistic settings. *arXiv preprint arXiv:2402.20245*.

Stefan Schweter. 2020. Italian bert and electra models.

Zhisong Zhang, Emma Strubell, and Eduard Hovy. 2022. A survey of active learning for natural language processing. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 6166–6190, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics.

Appendix A: Topic Modeling

In this section we present all the additional information about Topic Modeling.

Firstly, as mentioned above, we use gpt-4o-mini to label batches of tweets, and we elaborate the following prompt in order to stay under the api request per day rate limit:

```
"You are an expert Italian political content analyst. I will give you a batch of tweets from Italian politicians. For EACH tweet, determine if it contains significant political content or is just promotional content.
Significant tweets include:
• Policy positions or proposals
• Political criticism of opponents
• Commentary on current events
• Substantive discussions of issues
Non-significant (broadcast) tweets include:
• TV/radio appearance announcements
• Live stream announcements
• Schedule announcements
• Simple greetings without political substance
Respond with ONLY the tweet index number followed by either SIGNIFICANT or BROADCAST, one per line. Example format: 123: SIGNIFICANT; 124: BROADCAST; 125: SIGNIFICANT"
```

Successively, we expand on the interpretation of LDA MALLET which give us the following results for the right wing:

- ***National Values and Social Issues***
  This narrative centers on the protection of traditional "famiglia" (family) structures and "valori" (values), with frequent references to "vita" (life), "donna" (woman), "uomo" (man), and "preghiera" (prayer). Right-wing politicians emphasize "solidarietà" (solidarity) and "memoria"(memory) as pillars of Italian identity, constructing a narrative that places family at the nation's core.
  Care for vulnerable populations appears through discussions of "casa" (home), "mamma" (mother), "bimbo" (child), "papà" (father), "figlio" (son/child), "anziano" (elderly), and people with "disabilità" (disabilities), but notably within traditional family frameworks rather than through state welfare expansion. Concerns about "libertà" (freedom) and "diritti" (rights) are framed specifically in opposition to "islamico" (Islamic) extremism, "terrorismo" (terrorism), and "regime" (regime) threats like the "talebano" (Taliban), positioning right-wing politicians as defenders of Western civilization against external threats.
  A distinctive element is "anti-sinistra"

(anti-left) rhetoric, with frequent mentions of "minaccia" (threats), "violenza" (violence), and "insulto" (insults) allegedly coming from the left, countered with calls for "rispetto" (respect) and "fratello d'italia" (brotherhood of Italy) - also the name of a major right-wing party.
During the pandemic, right-wing discourse focused critically on "vaccino" (vaccine) policies, "scuola" (school) regulations, "covid" management, and government "misure" (measures). Words like "coprifuoco" (curfew), "chiusura" (closure), "pass" (health pass), and "greenpass" reveal opposition to restrictions, framed as threats to "libertà" (liberty) rather than public health necessities.

- ***Economic and Development Issues***
  The right-wing economic narrative champions "nucleare sicuro" (safe nuclear power) as a "pulito" (clean) and "energetico" (energetic) solution for Italy's "futuro" (future), alongside "gas" and "rinnovabile" (renewable) sources, positioning technological "ricerca" (research) as key to energy independence. During economic difficulties, right-wing politicians discuss "impresa" (business) struggles with "bolletta" (utility bills), "caro" (expensive) "luce" (electricity), and "gas" costs.
  Terms like "crisi" (crisis), "difficoltà" (difficulties), and "costo" (cost) frame economic problems, with "sostegno" (support) and "aiuto" (help) presented as solutions. On taxation, right-wing discourse criticizes "tassa" (tax) burdens on "impresa" (businesses) and champions "lavoratore" (worker) interests through "fiscale" (fiscal) reform.
  References to "cartella" (tax notices), "milione" (millions), "euro", "reddito" (income), and "pensione" (pension) reveal focus on tax relief and economic liberalization rather than government intervention. Regional development centers on "lavoro" (work) in both "sud" (south) and "nord" (north), with emphasis on "infrastruttura" (infrastructure), "ponte" (bridge) projects, and "risorsa" (resources). The "sviluppo" (development) narrative presents infrastructure as key to national unity and economic growth.

- ***Governance and International Relations***
  On international matters, right-wing politicians discuss "europeo" (European) affairs, "guerra" (war), "pace" (peace), and "difesa" (defense) through a nationalist lens, emphasizing "litalia" (Italy's) "paese" (country) interests first.
  The "Russia" - "Ucraina" (Ukraine) conflict is discussed with attention to economic and security implications for Italy. European relations use terms like "cooperation", "bilateral", "freedom", and "challenge" suggesting pragmatic engagement while preserving national sovereignty, with "energy" and "value" as recurring themes in international discussions.
  Immigration policy features prominently with "ministro" (minister) statements on "clandestino" (illegal immigrant), "confine" (border), "sbarco" (landings), and "immigrato" (immigrant).
  Words like "controllo" (control), "regola" (rules), and "illegale" (illegal) construct a security-focused narrative, with explicit criticism of "sinistra" (left-wing) immigration policies.
  Judicial matters focus on "giustizia" (justice) "riforma" (reform) through "referendum" campaigns, with "firma" (signatures), "gazebo" (information stands), and "raccolta" (collection) reflecting grassroots mobilization.
  The "processo" (legal process) is portrayed as needing conservative reform. Local governance highlights "sindaco" (mayor) races in "Roma", "Milano" and other regions, with "lega" (League party) "candidato" (candidates) and "territorio" (territory) concerns.
  The "squadra" (team) and "amico" (friend) language creates a personal, localized political narrative. Right-wing coalition building appears through "centrodestra" (center-right), "forza" (strength/Forza Italia party), and "liberale" (liberal) positioning.
  "Presidente" (president), "parlamento" (parliament), and references to opposition from the "sinistra" (left) and figures like "conte" and "salvini" complete the governance narrative.

Whereas for the left wing we get:

- ***International Relations and Progressive Values***
  This narrative weaves together concerns about the "guerra" (war) in "Ucraina" (Ukraine) and "Russia," with left-wing politicians advocating for "pace" (peace) while maintaining "sostegno" (support) through "diplomazia"

(diplomacy).
The humanitarian dimension focuses on "migranti" (migrants), addressing "tragedie" (tragedies) in the Mediterranean involving people fleeing from "Libia" (Libya), framed through appeals to "solidarietà" (solidarity) and "diritti umani" (human rights).
These international concerns connect to fundamental values of "libertà" (freedom), "democrazia" (democracy), and "diritti" (rights) within Italian society. The discourse highlights "donna" (women) and gender equality, often linked to historical "resistenza" (resistance) and anti-"mafia" struggles.
This creates a narrative connecting progressive politics with Italy's fight against authoritarianism.
The "europeo" (European) dimension positions Italy within a democratic "Europe," standing against authoritarian figures like "Putin" while advocating for collective action on shared challenges. Words like "valore" (value), "dovere" (duty), and "principio" (principle) permeate this discourse, framing political choices as moral imperatives.

- ***Economic and Social Policy***
This narrative addresses "crisi" (crisis) affecting "famiglie" (families) and "imprese" (businesses), particularly regarding "bollette" (utility bills), "costo" (costs), and "gas" prices. Left-wing politicians propose interventionist measures like "tetto" (price caps), government "intervento" (intervention), and "decreti" (decrees) to protect vulnerable citizens.
The "pandemia" (pandemic) has shaped discourse around "sanità" (healthcare), "vaccini" (vaccines), and balancing public health with "scuola" (education). This connects to arguments for strengthening "pubblico" (public) services and the welfare state.
"Lavoro" (work) and worker rights form a cornerstone, with calls for "salario minimo" (minimum wage), "reddito" (income) protections, and more progressive "tasse" (taxes) to address inequality. "Cultura" (culture), "musei" (museums), and "teatro" (theater) are framed not merely as heritage but as vital social assets. The narrative emphasizes future "investimenti" (investments) in renewable "energia" (energy), youth "opportunità" (opportunities), and "sociale" (social) programs through "PNRR" (National Recovery and Resilience Plan) and European funding.
Terms like "futuro" (future), "piano" (plan), and "programma" (program) signal a forward-looking approach to building a more equitable society.

- ***Governance and Political Positioning***
This narrative defines left-wing identity against the "destra" (right), with "antifascismo" (anti-fascism) serving as both historical reference and contemporary boundary. Left politicians oppose normalization of far-right ideologies in "stampa" (press), "scuola" (schools), and public discourse.
"Elezioni" (elections) feature prominently, with discussions of "candidati" (candidates), "coalizione" (coalition), and "primarie" (primaries). Left politicians criticize right-wing leaders like "Meloni," "Salvini," and "Berlusconi," often drawing parallels to international figures like "Trump".
The discourse includes reflections on "politica" (politics) itself, with the "sinistra" (left) positioning itself against "populismo" (populism). Terms like "serio" (serious) and "responsabile" (responsible) characterize the self-image of left parties as alternatives to right-wing governance.
"Roma" city management highlights issues like "rifiuti" (waste), "servizi" (services), and "quartieri" (neighborhoods), while broader "governo" (government) discussions address "parlamento" (parliament), "presidente" (president), and the "Quirinale" (presidential palace politics).
Throughout runs an aspirational tone about "cittadini" (citizens), "futuro" (future), and "importanza" (importance) of good governance, balancing critique with an affirmative vision using words like "grazie" (thank you), "bello" (beautiful), and "buono" (good) to convey optimism about left-wing leadership possibilities

Appendix B: Sentiment Analysis

Here we show a representative tweet for each tone label:

- neutral / informative: *"Fuori dal recinto. Intervista a Nicola Fratoianni"*
- positive / supportive / celebratory: *"Vitto-*

ria! Abbiamo votato al Parlamento europeo per difendere la produzione vinicola, un'eccellenza italiana portabandiera della nostra nazione nel mondo."

- angry / critical: *"Che sia incapacità o mancanza di volontà politica, Lamorgese ha fallito su tutto."*
- propaganda / calling to action: *"Ora tocca a te! #25settembrevotoLega"*

Appendix C: Results

C.1 Sentiment Evaluation

Class	Precision	Recall	F1-Score	Support
Propaganda	0.1828	0.0436	0.0704	390
Critical	0.6112	0.2967	0.3995	1102
Neutral	0.0966	0.6722	0.1690	180
Positive	0.1622	0.0245	0.0426	245
Accuracy			0.2457	1917
Macro Avg	0.2632	0.2593	0.1704	1917
Weighted Avg	0.4183	0.2457	0.2653	1917

Classification Report 350M Baseline

Class	Precision	Recall	F1-Score	Support
Propaganda	0.5582	0.8974	0.6883	390
Critical	0.9618	0.7541	0.8454	1102
Neutral	0.5146	0.4889	0.5014	180
Positive	0.7843	0.8163	0.8000	245
Accuracy			0.7663	1917
Macro Avg	0.7047	0.7392	0.7088	1917
Weighted Avg	0.8150	0.7663	0.7753	1917

Classification Report 350M Finetuned

Class	Precision	Recall	F1-Score	Support
Propaganda	0.2370	0.0821	0.1219	390
Critical	0.6608	0.4456	0.5322	1102
Neutral	0.0990	0.5500	0.1678	180
Positive	0.1282	0.0204	0.0352	245
Accuracy			0.3271	1917
Macro Avg	0.2813	0.2745	0.2143	1917
Weighted Avg	0.4538	0.3271	0.3510	1917

Classification Report 1B Baseline

Class	Precision	Recall	F1-Score	Support
Propaganda	0.5432	0.8872	0.6738	390
Critical	0.9234	0.6887	0.7890	1102
Neutral	0.3022	0.3778	0.3358	180
Positive	0.7983	0.7592	0.7782	245
Accuracy			0.7089	1917
Macro Avg	0.6418	0.6782	0.6442	1917
Weighted Avg	0.7717	0.7089	0.7216	1917

Classification Report 1B Finetuned

C.2 Topic Evaluation

Model	Accuracy	Precision	Recall	F1-Score
Baseline Right	0.428	0.390	0.428	0.397
Baseline Left	0.593	0.453	0.593	0.460
Finetuned Right	0.394	0.352	0.394	0.365
Finetuned Left	0.593	0.459	0.593	0.466
Baseline 1B Right	0.413	0.359	0.413	0.379

Topic Model Report