



POLITECNICO
MILANO 1863

**SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE**

EXECUTIVE SUMMARY OF THE THESIS

Leveraging Large Language Models in Italian Political Campaigns: Ad Generation, Similarity Analysis and ML Classification

LAUREA MAGISTRALE IN COMPUTER SCIENCE AND ENGINEERING - INGEGNERIA INFORMATICA

Author: ALESSANDRO MENCARELLI

Advisor: PROF. FRANCESCO PIERRI

Academic year: 2023-2024

1. Introduction

When in June 2017 eight scientists working at Google released “*Attention Is All You Need*”, a landmark research paper in machine learning introducing a new deep learning architecture known as the transformer – the groundwork of every Large Language Model (LLM) – it is not known if they were fully aware of the revolution they would cause in the years to come. Indeed, the advent of Large Language Models has marked a significant turning point in the landscape of artificial intelligence, fundamentally transforming how machines understand and generate human language. Leveraging huge datasets and complex neural network architectures, these models manage to perform innumerable tasks with unprecedented accuracy and fluidity, revolutionizing diverse fields by providing tools that augment human capabilities and streamline complex processes.

In the realm of political communication, LLMs are increasingly influencing how campaigns craft persuasive and targeted messages. Political campaigns have long relied on strategic messaging to influence public opinion, mobilize voters, and achieve electoral victories. The integration of LLMs offers the potential to enhance these strategies by generating tailored content that

targets specific demographics, analyzing voter sentiment, and optimizing message spread [1]. Despite the growing interest in leveraging LLMs for political purposes, existing research predominantly focuses on contexts outside Italy, lacking a study tailored to the Italian political landscape. The Italian political scenario is characterized by a distinct spectrum of political parties, coalition dynamics and regional macro-characteristics that significantly influence electoral strategies. Therefore, there is a pressing need for research that explores the capabilities of LLMs in generating persuasive political content suited to the specific demands and challenges of the Italian political arena.

To address this gap, this study undertakes a comprehensive analysis of the potential use of LLMs in Italian political campaigns. Leveraging data from the 2022 Italian general election – specifically Facebook political advertisements [4] – synthetic ads that reflect the ideological stances and communication styles of major Italian political parties will be generated. Employing state-of-the-art embedding techniques such as SBERT [5] and leveraging metrics like Cosine Similarity and BERTScore, the study conducts a detailed similarity analysis on both LLM-generated ads and their real coun-

terparts. Additionally, machine learning classifiers – based on Siamese Neural Networks, Random Forest and XGBoost – are developed and evaluated to determine their effectiveness in distinguishing partisan ideology and between authentic and synthetic ads. This dual approach not only demonstrates the potential of LLMs in engaging political communication but also establishes methodologies for monitoring and safeguarding the integrity of AI political advertising. By addressing the existing research gaps and proposing solid frameworks for analysis, this thesis aims to advance our understanding of how AI-driven technologies can be effectively and ethically integrated into political campaign strategies.

2. Background

The digital migration of political campaigns has brought online advertisements to the center of party political communication strategies. Studies have demonstrated that online political advertisements can effectively influence voter turnout and shape policy attitudes, particularly in competitive districts and among specific demographic groups. The persuasibility and impact of these ads, whether generic or microtargeted, has been the subject of extensive research within the research community.

2.1. Large Language Models and Persuasive Influence

Large Language Models (LLMs) have greatly stimulated the field of artificial intelligence, especially in the domain of natural language processing. These models are trained on very large datasets comprising text from the internet, books, and various media sources, learning to understand and generate human language. Essentially, they utilize deep learning architectures to model the statistical properties of language, allowing them to generate coherent text based on each request.

Studies have demonstrated the persuasive capabilities of LLMs in political discourse [1], highlighting their ability to craft tailored messages that resonate with specific demographics. Results show that LLM agents can simulate dynamics of persuasion by crafting arguments incorporating evidence-based facts showing trust and support. Such arguments were

ranked as most persuasive by humans and artificial agents alike, which entices the possibility of an influence by LLMs in opinion dynamics of online social media. Other experiments involved ChatGPT to compose personalized advertisements based on participants' known personality profiles showing how effectively LLMs can scale personalized persuasion [2]. The implications for election integrity are substantial. High-quality, AI-generated disinformation can be distributed rapidly and at scale, potentially influencing voter perceptions and undermining democratic processes. On the technical side, studies stating that the relationship between model size and persuasive effectiveness follows a log scaling law, indicating diminishing returns beyond a certain threshold [3].

2.2. Text Embedding and Similarity Metrics

In Natural Language Processing (NLP), a word embedding is a representation of a word. Typically, this representation is a vector that encodes the meaning of the word in such a way that the words that are closer in the vector space are expected to be similar in meaning. BERT is a transformer-based model that introduced a breakthrough in NLP by providing contextualized embeddings, where each word's representation depends on its surrounding context.

Once BERT embeddings have been crafted, the next step is to use them either directly for NLP models or as an object of similarity analysis. For the latter, the most widely used metric is cosine similarity, which is the cosine of the angle between the vectors calculated as the dot product of the vectors divided by the product of their lengths. In addition, some recent studies have also used as metric BERTScore, which uses BERT embeddings to score text similarities and is effective in judging the coherence and relevance of textual outputs.

2.3. The 2022 Italian General Election

Italian elections 2022 marked an important moment in recent Italian political history. Following the early dissolution of the Houses of Parliament, caused by Mario Draghi's resignation, elections were held on September 25, 2022. The main political parties and coalitions included:

- **Centre-right coalition:** Composed of Brothers of Italy (Fratelli d'Italia), the League (Lega), Forza Italia, and others.
- **Centre-left coalition:** Led by the Democratic Party (Partito Democratico) in alliance with smaller leftist and centrist parties.
- **Five Star Movement** (Movimento 5 Stelle).
- **Third Pole** (Terzo Polo): A centrist alliance between Azione and Italia Viva.

The election saw a shift to the right, with the Centre-right coalition winning a majority in both houses of parliament.

The 2022 Italian general election serves as an excellent case study due to its complex political landscape and the significant role of digital platforms in campaign strategies. Pierri's work [4] demonstrated that in this election, major political coalitions heavily invested in Facebook and Instagram advertisements, strategically targeting their traditional strongholds and certain gender/age demographics, thereby underscoring the critical role of digital platforms in contemporary electoral strategies.

3. Proposed Solution and Implementation

This thesis explores in a comprehensive way the potential use of large language models in a political election campaign, analyzing the various aspects related to it, starting with the generation of synthetic ads to assess how effectively LLMs can produce persuasive political ads that mirror the strategies of real-world campaigns, to end up testing if machine learning classifiers are capable of analyze the partisan origin and distinguishing between authentic and synthetic advertisements. The first step of the work was to obtain data on real political ads, from the 2022 Italian elections. Raw Facebook data, made available by Meta for research purposes, had been processed in [4] to label each ad with its corresponding political party. The final dataset, used in the similarity analysis, during the generation of the Ads with LLM and – in its unduplicated version – in the classifiers, consists of 963 real ads from 15 Facebook pages related to 7 parties.

3.1. Ad Generation Using LLMs

Meta-Llama-3.1-70B-Instruct, an open-source LLM, has been utilized for ad generation. Two prompting strategies were employed:

- **Few-Shot Learning:** Providing the model with a small number of examples before the main task.
- **Zero-Shot Learning:** The model generates responses without any prior examples or specific training data.

For the Few-Shot approach, ads were generated in three distinct batches by selecting top-spending real advertisements – from all pages, from official and party leader pages, and by grouping main parties – and then creating both generic and topic-specific ads based on these selections. In the Zero-Shot approach, summaries and bullet points of each party's political program, generated using **GPT-4o**, were provided. In order to enable proper analysis, the textual content of all advertisements – both real and generated ads – was then transformed through embedding techniques. Specifically, all the ads were embedded with the **SentenceTransformer** framework [5], using the specific model **nickprock/sentence-bert-base-italian-xxl-uncased**, tuned on Italian sentences.

3.2. Similarity Analysis

To understand the relationships and patterns among the generated and real political ads, a similarity analysis was conducted. The two metrics in measuring similarities used in the work are Cosine Similarity and BERTScore.

For each ad class and for each similarity metric an Heatmap had been generated to visualize and compare the similarity scores. These will intuitively lead to a view of how the ads within the same classes compare against one another, highlighting large clusters or outliers. Additionally, both intra-party and inter-party similarities were analyzed to understand how ads within the same party compared to those from different parties. Intra-party similarity focused on the consistency of messaging within a single party, while inter-party similarity refers to the distance of one party compared to all the others.

To further explore the structure of the datasets, for each of the aforementioned ad classes, it were tested two clustering techniques: t-SNE and K-Means. These clustering analyses, combined

with the results of the similarity analysis, were aimed at assessing possible positive use by ML classifiers, and an estimate on the ad classes and tasks in which they could perform best.

3.3. Machine Learning Classifiers

The final component of the thesis involves the development and evaluation of machine learning classifiers designed to perform two key tasks among the different data sources that have been created so far: predicting the political party associated with a given ad and distinguishing between real and generated advertisements. These models were trained and evaluated using stratified cross-validation to ensure reliable and balanced performance across parties. Performance metrics including Accuracy, F1-score, and AUC are used to evaluate the effectiveness of each model, providing a comprehensive understanding of their strengths and limitations.

Three types of classifiers are employed:

Siamese Neural Networks The architecture consisted of a custom three-layer Siamese network trained using contrastive loss, in which each ad in the pair – via its embedding representation – were passed through identical sub-networks that shared weights. The contrastive loss function aims at minimizing the difference between embeddings from ads of the same party by maximizing the embedding distance of ads from different parties. The dataset used were: **real-ads**, **real-ads-cap**, **few-shots**, and **zero-shot**; a different SNN was trained for each dataset.

PartyClassifiers PartyClassifiers were designed to predict the political party associated with a given ad testing two type of classifier: **RandomForestClassifier** from **sklearn** library and **XGBClassifier** from **xgboost**.

The model outputs its prediction, i.e. the name of the party, and, if desired, its percentage of sureness.

VsClassifiers VsClassifiers were designed for a distinct task, specifically to distinguish between real and generated advertisements. Similar to the PartyClassifiers, the two base models that has been employed are Random Forest and XGBoost Classifier. On the contrary, the



Figure 1: Few-Shot Generated Ads Similarity Heatmap (Cosine Similarity).

datasets differ, with each containing two distinct classes: **real-vs-fs** with 940 ads (470 real and 470 few-shot generated), **real-vs-zs** with 984 ads (492 real and 492 zero-shot generated), and **real-vs-gen** with 1,270 ads (635 real and 635 generated).

The model outputs its prediction – this time binary: *Real/Generated* – and, if desired, its percentage of sureness.

4. Results

4.1. Similarity Analysis

The similarity analysis revealed that generated ads exhibit high internal similarity within each party and maintain distinctiveness across different parties. The heatmaps for real and generated ads showed higher similarity scores along the diagonal, indicating greater likeness within the same party. Interestingly, taking a party, the synthetic ads tend to be more similar to each other than to the ads used as few-shot examples to generate them. This points to an ability of LLMs to abstract from the mere examples provided and create a message based on the predicted ideology of the party in question.

The heatmaps for Zero-Shot generated ads show uniformly higher similarity scores across all classes. This behavior suggests that in this setting the LLM is less effective in differentiating ads regardless of their party affiliation, generating more generic and similar ads. Despite this, the distinction between intra-party and inter-party similarities remains clear and it is even

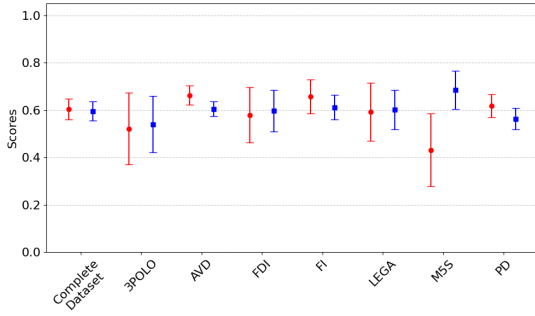


Figure 2: Siamese NN Performance on Real-Ads-cap (red: f1-score, blue: accuracy).

more pronounced compared to the Few-Shot scenario, still allowing ads to be effectively grouped by party.

The clustering results emerged in line with the findings from the similarity analysis section. In fact, generated ads, both Few-Shot and Zero-Shot showed a greater inclination to form distinct and easily identifiable clusters with respect to Real Ads.

4.2. Ad Classification

4.2.1 Siamese Neural Networks

The Siamese Neural Network demonstrated decent performance on the Real-Ads dataset. Due to the imbalance in the dataset, when trained on initial real-ads dataset certain subclasses exhibited poor results. However, performance significantly stabilized and improved for underrepresented parties when trained using the *real-ads-cap* version, balancing ads per party. When trained on generated ads, the SNN achieved high F1-scores and accuracy across all parties, indicating effectiveness in identifying and classifying LLM-generated content.

4.2.2 PartyClassifiers

As previously, they demonstrated decent performing when trained on Real-Ads dataset in his *complete* version, especially XGBoost-based version. When trained with the *cap* version, performance slightly improved and balanced across the different parties, especially for the RandomForest-based models. Regarding LLM generated ads both classifiers exhibit high F1-scores across all parties. Relative to party subclasses, some achieve higher results than others, likely due to the greater distinctiveness of lan-

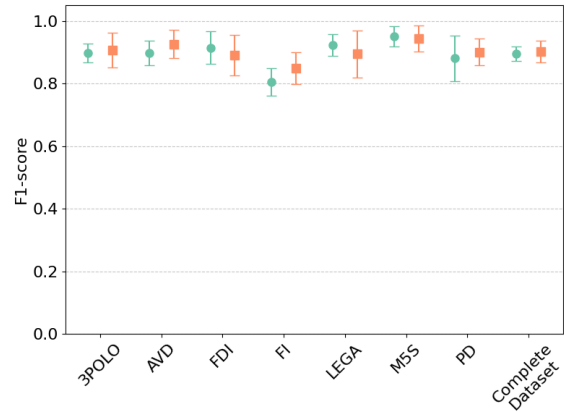


Figure 3: PartyClassifier F1-score Performance on Few-Shot Ads (green: RF, orange: XGB).

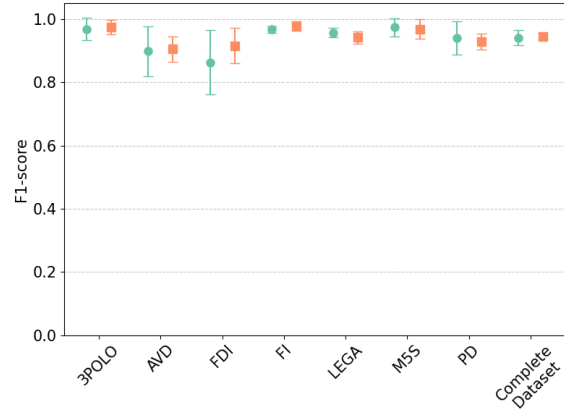


Figure 4: VsClassifier F1-score Performance on Real vs. Gen Ads (green: RF, orange: XGB).

guage in certain parties' campaigns.

4.2.3 VsClassifiers

The results indicate that the classifiers performed exceptionally well across all scenarios, achieving high F1-scores in distinguishing between real and generated ads. In addition, analysis of the party subclasses shows that the models behave consistently, demonstrating their ability to extract patterns that are valid across all parties rather than being confined to specific topics or language styles associated with only some of them.

4.3. Human Evaluation

To further validate the performance of the machine learning classifiers, a simple test experiment involving human evaluation was conducted. The results indicate that human judg-

ment is only slightly more skillful than flipping a coin, with a $F1\text{-Score}_{\text{Mean}} = 0.45 \pm 0.16$ for human evaluators. This stark contrast underscores two key concepts: the reliability and efficiency of machine learning models in classification tasks compared to human judgment, and the difficulty of the human in determining whether an ad is AI-Generated, confirming the effectiveness of using LLM in the policy arena.

5. Conclusion

This thesis has comprehensively examined the meeting points between large language models and political communication, with a specific focus on their application in generating political advertisements using the 2022 Italian general election as a case study.

The study demonstrated that LLMs, specifically **Llama 3.1**, can effectively replicate the persuasive and thematic elements of authentic political advertisements, maintaining high similarity within parties and distinctiveness between them through similarity metrics and clustering analyses, thereby showing their capability to craft targeted and ideologically aligned campaign content akin to actual political strategies.

The development and evaluation of ML classifiers further validated the effectiveness of the generated ads by achieving high accuracy in predicting party affiliations, thereby confirming that LLMs are capable of generating precisely targeted and tailored advertisements for specific political parties. Secondly, the classifiers demonstrated a superior ability to distinguish real from generated ads compared to humans, highlighting the potential of machine learning for automating political content analysis and safeguarding the integrity of political milieu.

These results bring to light that the use of LLM in the political scenario – including the Italian one – is already realistic, and it is able to both augment and challenge traditional political communication strategies. Considering the truthfulness that synthetic ads and the seemingly endless potential for personalization and targeting that Large Language Models offer, we are in front of a radical change in the way we have been analyzing political campaigns so far. De facto, this ensures that anyone, with meager expense and required know-how, can create a professional and widely diversified political ad-

vertising strategy.

The ability of LLMs to produce highly convincing and targeted content undermines traditional safeguards by making powerful capabilities widely accessible, raising significant ethical concerns regarding their misuse in political advertising and necessitating responsible development and regulation.

Starting from the limitations of this work, future research should explore direct testing of LLMs on classification tasks and conduct real-world experiments involving human participants to assess the truthfulness and persuasiveness of AI-generated ads.

Acknowledgements

I would like to thank my supervisor, my friends, my fiancé and my family for all the support, love, and precious advice that I received.

*Impossible is not a fact, it's an opinion, so go
build the future.*

LOGAN KILPATRICK

References

- [1] S. M. Breum, D. V. Egdal, V. G. Mortensen, A. G. Møller, and L. M. Aiello. The persuasive power of large language models. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 18, pages 152–163, 2024.
- [2] K. Hackenburg and H. Margetts. Evaluating the persuasive influence of political microtargeting with large language models. *Proceedings of the National Academy of Sciences*, 121(24):e2403116121, 2024.
- [3] K. Hackenburg, B. M. Tappin, P. Röttger, S. Hale, J. Bright, and H. Margetts. Evidence of a log scaling law for political persuasion with large language models. *arXiv preprint arXiv:2406.14508*, 2024.
- [4] F. Pierri. Political advertisement on facebook and instagram in the run up to 2022 italian general election. In *Proceedings of the 15th ACM Web Science Conference 2023*, pages 13–22, 2023.
- [5] N. Reimers. Sentence-bert: Sentence embeddings using siamese bert-networks. *arXiv preprint arXiv:1908.10084*, 2019.