

## Painting the Senate #Green: A Corpus Study of Twitter Sentiment Towards the Italian Environmentalist Blitz<sup>1</sup>

**Antonio Bianco**  
University of  
Bergamo/Pavia  
Piazza del Lino, 2, 27100  
Pavia (PV), Italy  
antonio.bianco@unibg.it

**Claudia Roberta Combei**  
University of Pavia  
Piazza del Lino, 2, 27100  
Pavia (PV), Italy  
claudiaroberta.combei@unipv.it

**Chiara Zanchi**  
University of Pavia  
Piazza del Lino, 2, 27100  
Pavia (PV), Italy  
chiara.zanchi01@unipv.it

### Abstract

This study analyzes the reactions of the Italian Twitter community to an environmental demonstration that occurred in Rome on January 2<sup>nd</sup>, 2023. We compiled a corpus of 368,531 tokens consisting of 11,780 tweets, collected during a 7-day period. We propose a mixed-method approach that combines automated and manual corpus analyses of sentiment, emotions, and implicit language. Our findings offer insights into how tweets reflected the users' attitudes toward a variety of subjects and entities. Although the sentiment of the overall debate was distributed rather evenly, the incident itself seems to have sparked negative sentiment and emotions among Twitter users. The results of our manual analyses revealed some issues with respect to the automatic classification of sentiment, due to the fact that some tweets contained irony, sarcasm, and slurs. Non-literal interpretations were ignored by the tools at hand that could not account for complex rhetorical-argumentative strategies.

**Keywords:** Twitter discourse; Italian; sentiment analysis; environmental demonstration

**DOI:** 10.28995/2075-7182-2023-22-1021-1031

## Раскрашивание Сената в зелёный цвет. Корпусный анализ тональности в Твиттере по поводу акции итальянских экологов

**Антонио Бьянко**  
Университет  
Бергамо/Павии  
Пьяцца дель Лино, 2,  
27100 Павия (PV), Италия  
antonio.bianco@unibg.it

**Клаудия Роберта Комбей**  
Университет Павии  
Пьяцца дель Лино, 2, 27100  
Павия (PV), Италия  
claudiaroberta.combei@unipv.it

**Кьяра Занки**  
Университет Павии  
Пьяцца дель Лино, 2, 27100  
Павия (PV), Италия  
chiara.zanchi01@unipv.it

### Аннотация

В данном исследовании анализируется реакция итальянского сообщества в Твиттере на экологическую демонстрацию, которая прошла в Риме 2 января 2023 года. Для этой цели мы составили корпус из 368531 токена, включающий 11780 твитов, собранных в течение 7 дней. Нами был использован смешанный метод, сочетающий автоматический и ручной анализ тональности, эмоций и имплицитности языка на материале корпуса. Результаты показывают, что твиты отражают отношение пользователей к различным субъектам и организациям. Хотя тональность в ходе общих дебатов распределилась довольно равномерно, сам инцидент вызвал среди пользователей Твиттера негативные тональность и эмоции. Результаты ручного анализа выявили некоторые спорные случаи автоматической классификации тональности, связанные с тем, что некоторые твиты содержали иронию, сарказм и оскорбления. Поскольку использованные инструменты не учитывали сложные риторико-аргументативные стратегии, небуквальные интерпретации игнорировались.

**Ключевые слова:** дискурс в Твиттере; итальянский; анализ тональности текста; экологическая демонстрация

---

<sup>1</sup> This work represents a collaborative effort undertaken by all three authors, who have consistently contributed throughout its development. In accordance with the academic requirements in Italy, authorship attribution is specified as follows: Antonio Bianco authored section 4, Claudia Roberta Combei authored sections 2 and 3, and Chiara Zanchi authored sections 1 and 5.

## 1 Introduction

This paper analyzes the Italian Twitter users’ reactions toward one of the latest “civil disobedience actions” by a group of environmental activists called *Ultima Generazione* ‘last generation’ (<https://ultima-generazione.com/chi-siamo/>). On January 2<sup>nd</sup>, 2023, five young activists splattered orange washable paint onto the façade of the Italian Senate to protest the immobilism of Italian politics at facing the current “ecoclimatic collapse”.<sup>2</sup> This demonstration has been both minimized and strongly condemned; for sure, it raised a fierce debate in the Italian public opinion.

The paper unfolds as follows. Section 2 describes how our corpus of tweets was collected and preprocessed, as well as the tools used to analyze it. Section 3 contains the quantitative analyses. Section 4 qualitatively comments upon tweets that reveal to be challenging for sentiment analysis. Section 5 concludes the paper.

## 2 Data and methods

Over the last 20 years, there has been a significant change in how people interact. The development of social media platforms, such as Twitter or Facebook has played a crucial role in this transformation. Today, people use Twitter to stay informed, build networks, establish ingroup identities, and express opinions. Twitter has been, in fact, the subject of numerous studies that employed sentiment analysis tools to measure the attitudes of the public opinion on a variety of topics ([1], [2], [3]).

Our paper aims to measure the opinions of the Italian Twitter community regarding the environmental blitz described in section 1. For this goal, we collected a sample of 53,301 tweets using the Twitter API and the `academicwitteR` package [4] for R [5]. The complete dataset consisted of all Italian tweets, retweets, and quotes that matched our search criteria, based on a set of hashtags related to the event, shown below in Table 1:

Hashtag	English translation	Hashtag	English translation
#alessandrosulis	#alessandrosulis	#ecovandali	#ecovandals
#blitz	#blitz	#gas	#gas
#carabinieri	#carabinieri	#imbrattato	#stained
#clima	#climate	#inquinamento	#pollution
#climateemergency	#climateemergency	#larussa	#larussa
#crisiclimatica	#climatecrisis	#lauraparacini	#lauraparacini
#davidenensi	#davidenensi	#senato	#senate
#digos	#digos	#ultimagenerazione	#lastgeneration
#ecocretini	#ecocretins		

Table 1: The hashtags used for the queries

Even though the episode happened on January 2<sup>nd</sup>, 2023, at 08:00 AM, we extended our period of analysis to a full week, from January 2<sup>nd</sup> 5:00 AM through January 9<sup>th</sup> at 12:00 AM. This allowed us to also capture tweets written the days after the incident that might have contributed to framing the debate on this topic across Twitter. All duplicates in the corpus were automatically removed and the remaining data were manually examined to eliminate irrelevant tweets. The resulting corpus included 11,780 original tweets, consisting of 368,531 tokens, for a total of 23,733 lemmas. The tweets were processed and analyzed, both automatically and manually, to account for the perceptions and attitudes of the users.

The computerized corpus investigation consisted of several analyses. First, we explored the distribution of the most frequent lemmas in this debate. The corpus was lemmatized using the pre-trained model UD Italian PoSTWITA ([6], [7]) for tweets, available in the `udpipe` package [8] for R. The lemmatized corpus was manually adjusted to correct recurrent issues, such as the unnecessary lemmatization of proper names and the incorrect lemmatization of abbreviations, keywords, and

<sup>2</sup> The news can be read here: <https://video.corriere.it/cronaca/ultimo-blitz-ambientalisti-imbrattano-facciata-senato-fermati/21cb0ef6-8a71-11ed-8b19-cdc718310dd5>

misspelled text. The first 40 most frequent lemmas were plotted using the `ggplot` function from the `tidyverse` package [9] for R.

Given the main aim of this study, we employed a mixed-method approach to measure sentiment and emotions. We started from a lexicon-based technique that allowed us to map the lemmatized tweets to the Italian version of the `nrc` resource for sentiment and emotions proposed by [10] and made available through the packages `syuzhet` [11] and `tidytext` [12] for R. This readily available 13,875-word lexicon is based on [13] and [14], and identifies eight primary emotions: *trust*, *joy*, *anticipation*, *surprise*, *sadness*, *disgust*, *anger*, and *fear*. The approach assumes that a word may be associated with one or more emotions. If a lemma in the corpus is not matched to any lexicon entry, it will not be assigned sentiment and emotion scores. Before using this lexicon-based tool, we applied three stop-words lists for Italian (`snowball`, `stopwords-iso`, and `nlk`), available on the `tidytext` package, to remove lexically empty or uninformative words (e.g., prepositions, conjunctions, etc.). Numbers, punctuation, and one-character sequences were also stripped from the corpus. Sentiment is analyzed with the `get_sentiment`, `rescale_x_2`, and `get_dct_transform` functions in the `syuzhet` package that iterate over each tweet and return numeric values (from -1 for extremely negative to 1 for extremely positive) based on the mapping with the `nrc` lexicon. In addition, the `get_sentiment` function assigns numerical values to each primary emotion detected in the tweet.

The other technique proposed here regards the sentiment analysis of tweets with the tools of the `TextBlob` library [15] for Python, built upon the `NLTK` library [16]. Before analyzing the corpus, we automatically translated the tweets into English with the `Googletrans` library [17]. The potential errors generated by the translation process are a limitation of this study. We are aware that some subtleties of the original tweets were lost or altered in translation (e.g., the word *gretino*, a pun combining *cretino* ‘cretin’ and *Greta*, a reference to the environmental activist Greta Thunberg, was mistranslated as ‘Greek’). Nevertheless, we think that the solution proposed here, which combines machine translation and `TextBlob`, an established approach in the scholarly research on sentiment analysis in languages other than English [18], can be useful in assessing the Twitter users’ perceptions of this incident. The `TextBlob` library employs a Naïve Bayes classifier pre-trained on film reviews and considers the weight of intensifiers, emoticons, emojis, and exclamation marks to calculate the sentiment. The score is expressed on a scale between -1 (extremely negative) to 1 (extremely positive). Using average sentiment scores may help summarize and simplify sentiment analysis results, providing a concise representation of the overall sentiment in a corpus. Nevertheless, it is important to acknowledge that averages should not be regarded as anything more than what they are: measures of central tendency that fail to capture the intricacies or context-specific sentiment. Therefore, the average scores at the corpus level for both systems are calculated for comparison purposes only.

Parallely, a sample<sup>3</sup> of 49 tweets was built for an experiment with ChatGPT, to test how this advanced language model interpreted the sentiment of some corpus data and to compare its performance to our NLP systems. The qualitative analyses examined the use of irony within this sample of tweets. This specific focus was chosen due to the inherent limitations of automated sentiment detection tools, particularly in relation to ironic tweets, as shown in sections 3 and 4. By conducting manual analyses, we also delved into the persuasive impact of slurs within the said set of tweets.

The results of the analyses were plotted with the `Matplotlib` visualization library for Python [19].

### 3 Quantitative analyses

Extracting the most frequent lemmas allowed us to gain preliminary insights into the subtopics discussed in the debate around the incident described in section 1. We intentionally retained words closely associated with the environmental demonstration under investigation, as our objective was to comprehend the significance of each subtopic within our corpus. As Figure 1 shows, *senato* ‘senate’ was the most frequent word in the corpus, holding a prominent position and appearing frequently both as a reference to the institution and the building. The second and the third most frequent lemmas were *imbrattare* ‘to stain’ and *vernice* ‘paint’; the choice of *imbrattare* may suggest that painting the senate walls (*muro* ‘wall’ and *facciata* ‘façade’, other frequent lemmas) was largely perceived as a vandalism act on the Twitter arena. The high frequency of the hashtag *#ultimagerazione* ‘last generation’ and of

<sup>3</sup> The sample is here: [https://osf.io/5jve9/?view\\_only=779b0287c774498ea85ea1b5d507ea27](https://osf.io/5jve9/?view_only=779b0287c774498ea85ea1b5d507ea27)

the lemmas *ultimo* ‘last’, *generazione* ‘generation’, *attivista* ‘activist’, *ambientalista* ‘environmentalist’, *ragazzo* ‘boy’, and *ragazza* ‘girl’ indicates that Twitter users identified the environmentalist activists with the younger generation. The high frequency of lemmas such as *clima* ‘climate’, *climatico* ‘climatic’, *ambiente* ‘environment’, placed at the center of the discussion the environmental motivation of the protest. However, environmental issues were viewed as either a valid or an inadequate justification of the incident, depending on the users’ perspective. Lemmas such as *fascista* ‘fascist’ and *civile* ‘civil’ permeated the discourse, implying that Twitter users discussed the appropriate ways of protesting in a democratic society. This episode received, in fact, attention from politicians, especially from the president of the senate, Ignazio La Russa (his name is frequent in the corpus), imputing potential political implications.

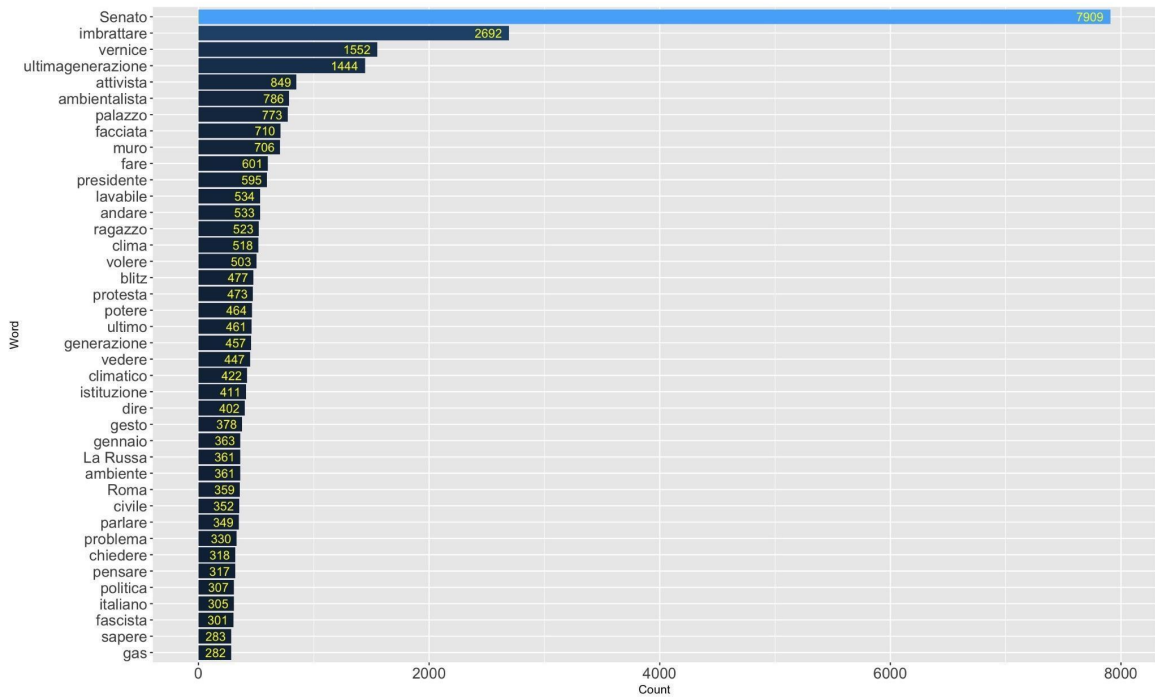


Figure 1: Top-40 most frequent lemmas

Next, we used the nrc lexicon to measure the sentiment and emotions conveyed in the debate. The overall average sentiment during the week considered was -0.059 (sd: 0.63), while the percentage of tweets labeled as negative, neutral, or positive is shown in Figure 2.

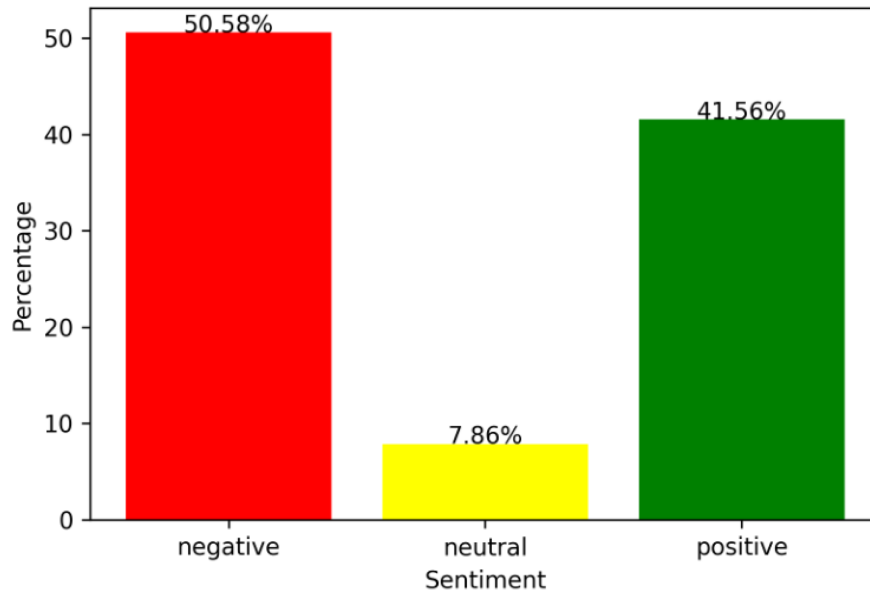


Figure 2: The distribution of sentiment based on nrc system

The average sentiment was only slightly negative (still, -0.059 is very close to a neutral score), but the fact that over 50% of the tweets were classified as negative may suggest that this environmental blitz was not debated in positive terms. However, the standard deviation of 0.63 indicates that tweets conveyed a wide range of sentiments: there was considerable variation in the users’ attitudes with respect to this episode. For a better understanding of these results, we measured the emotions transmitted by the tweets. Figure 3 displays the percentage of words that were associated with each of Plutchik’s primary emotions.

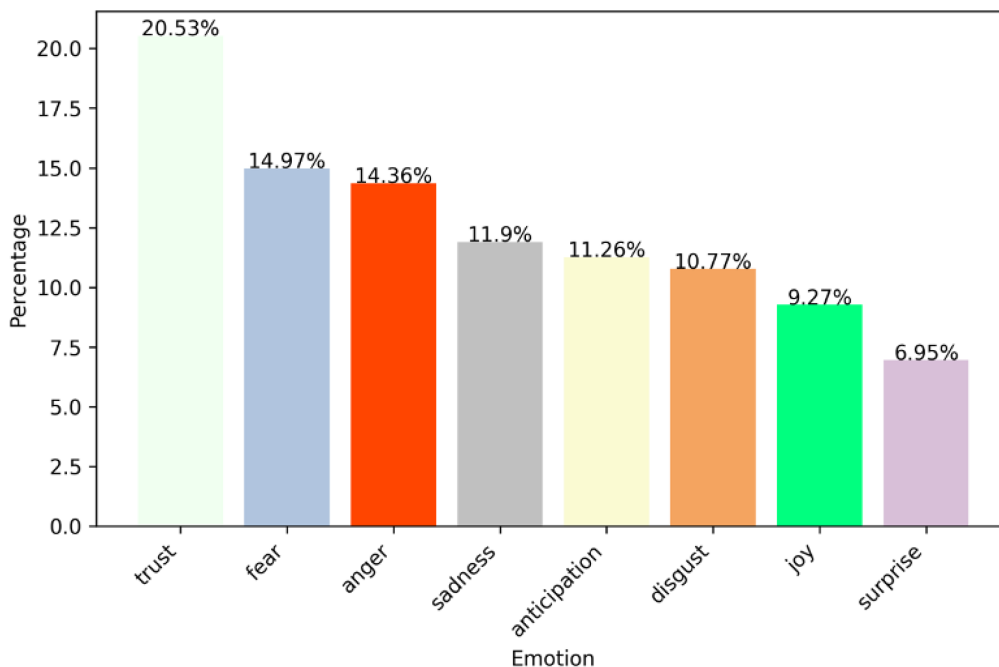


Figure 3: The distribution of emotions based on nrc system

Trust was the most frequent emotion, probably due to the frequent occurrence of political and institutional bodies and names (*senato* ‘senate’, *presidente* ‘president’, *politica* ‘politics’, etc.). Nevertheless, words related to trust did not always imply positive sentiment or support for politics and institutions; rather, they could transmit distrust or skepticism. In general, the incident appeared to have

elicited strong reactions among Twitter users, with a relatively higher occurrence of words associated with negative emotions, such as fear and anger. Specifically, words related to anger (*idiota* ‘idiot’, *delinquente* ‘delinquent’, *vergognoso* ‘shameful’, *condannare* ‘condemn’, etc.) seemed to be linked to a sense of breach of civil norms, while those evoking fear (*paura* ‘fear’, *attacco* ‘attack’, *pericoloso* ‘dangerous’, *danno* ‘damage’, etc.) tended to reflect worries about the possible effects of the incident. Finally, the episode did not appear to have surprised Twitter users, given the low percentage of words associated with this emotion.

To better capture the sentiment of the Twitter debate, we also measured it with TextBlob. The average sentiment score during the week considered was 0.035 (sd: 0.26), that appears aligned to the score we reported for the nrc system. However, at this point, it is important to note that when sentiment is extracted from a large corpus of tweets, its average value at corpus level tends to converge towards 0, since positive and negative values cancel one another out. This does not necessarily mean that there is a bigger percentage of neutral sentiment. Rather, it implies that the average score at the corpus level approaches 0 due to the counterbalancing effect of positive and negative sentiment expressed within the tweets.

For this reason, we also explored the distribution of sentiment as a function of time, as displayed in Figure 4.

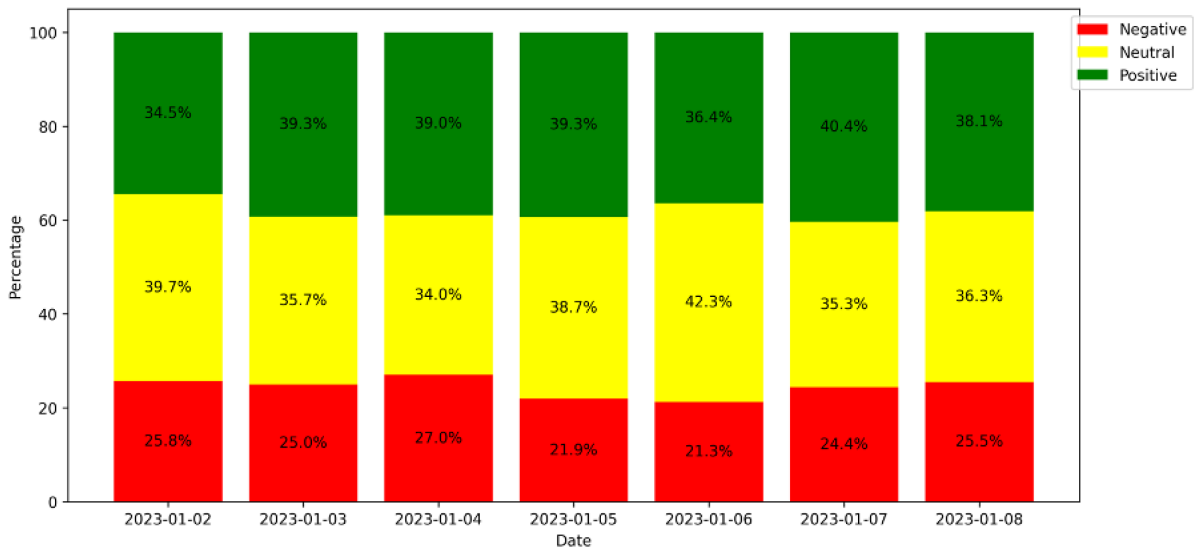


Figure 4: The percentage of sentiment as a function of time based on TextBlob system

In contrast to the findings of the nrc system, which exhibited a significant predominance of either negative (50.58%) or positive (41.56%) sentiment values, accompanied by a relatively low percentage of neutral tweets (7.86%), Figure 4 depicts a more balanced distribution across the three classes as regards the results obtained with the TextBlob system. In fact, 38.06% of the tweets were classified as positive, 36.68% as neutral, and 25.26% as negative. The largest proportion of tweets with a negative sentiment were published on January 4<sup>th</sup> - two days after the event - while most positive tweets occurred on January 7<sup>th</sup>. Instead, January 6<sup>th</sup> had the highest percentage of neutral tweets. Figure 4 suggests that the perspectives and responses of public opinion towards the incident changed only slightly during the week considered, with the results of the repeated measures ANOVA indicating that there were no statistically significant differences in the distribution of the three classes of sentiment over the seven days (F-value = 0.00187, p-value > 0.05,  $\eta^2 < 0.001$ ). At the same time, the value of the standard deviation (0.26) at the corpus level suggests that the reactions of the Twitter users were not unanimous.

Due to the large number of tweets analyzed automatically, it was difficult to verify manually the accuracy of each label as well as to match it to the intended targets of the sentiment and emotions transmitted. In fact, tweets addressed several topics related to the environmentalist episode, and the sentiment reflected the users’ perception and reactions towards a variety of subjects and entities, including the environmental activists of Last Generation, the incident itself, the Senate, several politicians, and governmental organizations such as *Carabinieri*. Therefore, we purposely selected a

sample of 49 tweets that posed challenges to the sentiment analysis tools we used, because they contained irony, sarcasm, or slurs. In fact, these linguistic devices significantly impacted the sentiment of the tweets, and in many cases, the nrc system (and to a lesser extent, TextBlob) failed to accurately identify the correct sentiment score of tweets conveying irony and sarcasm. Before conducting the qualitative manual analysis of these data (see section 4), we also wanted to determine whether sophisticated language models, such as the one used for ChatGPT, could assign the proper sentiment and emotion labels to this type of challenging tweets. The said sample of 49 tweets was classified by ChatGPT into the following categories: 40 tweets as negative, 8 as neutral, and 1 as positive. In general, the labels reflected the sentiment we associated with these tweets, indicating rather correct identifications of instances of irony, sarcasm, amusement, and expressions evoking disgust and resentment. Nevertheless, 4 tweets with ironic content expressing negative sentiment towards the activists and the act of staining the walls were misclassified by ChatGPT as neutral instead of negative. Some of these tweets are commented on in section 4 below, specifically tweets (6)-(7).

## 4 Qualitative analyses

This section discusses irony and derogatory epithets or slurs, which are essential for understanding the Twitter debate on the environmental blitz, but which posed challenges to our NLP tools for sentiment analysis. Ignoring these phenomena may lead to a misinterpretation of what lies beneath the explicit content of tweets.

### 4.1 The interplay between irony and sentiment analysis

Figurative language (e.g., irony, metaphors, hyperboles) is one of the most complex aspects of human language, when it comes to NLP sentiment analysis ([20], [21]). Irony deserves particular attention, as it significantly characterizes the Twitter media ecosystem [20].

Irony is an implicit linguistic device: a content textually unavailable but nonetheless conveyed and present in the locutor's communicative intentions [22]. In fact, it is a conversational implicature generally induced by an intentional violation of the Maxim of Quality [23]. With an ironic statement, speakers usually want to communicate the opposite of what is explicitly said [24]. Thus, ironic utterances may reverse the polarity of the message in which they occur [25]. Since irony is contextually dependent and its decoding requires complex mental operations, NLP tools may fail in classifying the sentiment of ironic statements.

In fact, we encountered three main issues when using the nrc lexicon and the TextBlob system for sentiment analysis. First, some tweets, classified as having a negative sentiment toward the Last Generation's act, actually sympathized with activists (1)-(3). Second, some tweets that apparently expressed positive sentiment toward the activists' actions in fact ironically criticized it (4)-(6). Third, some tweets (7)-(8) that were assigned neutral and/or positive sentiment, due to irony, conveyed negative sentiments toward the activists (7) or mocked the incident (8).

Since irony recognition is largely dependent on the reader's subjectivity [26], we resorted to a qualitative analysis to investigate the sentiment of this type of tweets.

	<b>Original tweet</b>	<b>Translation into English</b>	<b>nrc polarity</b>	<b>TextBlob polarity</b>
(1)	Ma parliamo un altro po' di quella vernice brutta brutta lanciata sul Senato!	But let's talk a bit more about that very nasty paint being thrown over the Senate!	-0.86	-0.27
(2)	#evasorifiscali in libertà e dei ragazzi per 2 graffiti con vernice lavabile in carcere	#taxevaders free and some boys for 2 graffiti with washable paint in prison	-0.22	-0.35
(3)	Stiamo vivendo una crisi climatica irreversibile ma povero muro del senato	We are living in an irreversible climate crisis but poor senate wall	-1	-0.4

Tweets in (1)-(3) imply ironically the opposite of what is explicitly communicated, minimizing the activists' actions and showing solidarity with them. In addition, the opposite of (1) conveys a sort of impatience regarding the prominence this news was given in the media. In (1), irony is signaled both by the exclamation point and the reduplication of *brutta* 'nasty'; this adjective led our tools to classify this tweet as negative. Reduplication, intensifying the adjective meaning, makes the literal interpretation of (1) too overstated to be taken as true (also, adjective reduplication is one of the features of Italian baby talk). The occurrence of a term expressing judgments and feelings (*brutta* 'nasty') is a further indication of irony [26]. In (2) mentioning tax evaders' freedom is unexpected, and such freedom clashes with the fact that young activists are in prison, despite using washable (and not permanent) paint. It is precisely such unexpectedness [27] that allows us to infer the opposite of what is stated. Similar considerations concern the tweet shown in (3). Indeed, based on the author's viewpoint, it seems senseless to engage in condemning a protest action while ignoring irreversible climate change.

	<b>Original tweet</b>	<b>Translation into English</b>	<b>nrc polarity</b>	<b>TextBlob polarity</b>
(4)	Buongiorno dal #Senato. Anche oggi abbiamo salvato il clima!	Good morning from the #Senate. Even today we saved the climate!	0.6	0.7
(5)	Con il Senato imbrattato, il clima sarà sicuramente migliore	With the Senate smeared, the climate will certainly be better	1.0	0.10
(6)	In realtà sono degli artisti lasciate che sfoghino liberamente la loro creatività	In fact they are artists let them freely vent their creativity	0.1	0.2

Tweets in (4)-(6) were classified as positive by both systems. The overstatements of (4) and (5) lead us to interpret these tweets as ironic. The same applies to (6). What is said in (6) seems to be a violation of the Maxim of Quality, as the locutor cannot *bona fide* assert that what happened was an artistic gesture. Thus (4)-(6) convey dissent towards Last Generation's action.



	<b>Original tweet</b>	<b>Translation into English</b>	<b>nrc polarity</b>	<b>TextBlob polarity</b>
(7)	ultima generazione lo spero proprio	last generation I really hope so	0.6	0.0
(8)	Coi secchi di vernice coloriamo tutti i muri, case, vicoli e palazzi, perché lei ama i colori. #Senato	With buckets of paint let's color all the walls, houses, alleys and buildings, because she loves colors. #Senate	0.0	0.0

Tweets (7)-(8) exemplify particular types of irony. Specifically, (7) is a case of sarcasm, a sharper kind of irony. To understand (7), one does not need to infer the opposite of what is asserted; irony is triggered by joking with the ambiguity [21] of *ultima generazione* 'last generation'. The expression is not used to properly name an environmentalist group but is to be interpreted literally. Therefore, the tweet is to be read as 'I hope this is the last generation of (this type) of activists'. So, the sentiment assigned by the tools is inconsistent with the sentiment that arises from what is implicitly communicated (ChatGPT also misclassified this tweet as neutral, see section 3).

In (8) a case of echoic irony [28] is exemplified. The Last Generation's act is described by echoing words of a well-known song by the singer Riccardo Cocciante. The hashtag #Senate refers to the incident. The irony arises as the motivations underlying the event described in the song (a romantic gesture) and the activists' actions are completely different. In (8), irony generates hilarious effects that tend to mitigate the disagreement.

#### 4.2 The power of slurs

In many tweets, including (9)-(12), Last Generation activists were designated by slurs [29]. Slurs have a derogatory connotation that eases the recognition of a negative sentiment by the systems we used. The (persuasive) appeal of slurs goes beyond their explicit derogatory connotation, as these expressions convey implicit meanings [30]. In particular, slurs allow the locutors to implicitly evoke stereotypes associated with the criticized social groups, in this case Last Generation activists, and to indirectly communicate the superiority of their own group. As a result, slurs, through their association with stereotypes, have the power to (implicitly) elicit a greater degree of negative sentiment compared to other negative words. In addition, stereotypes pertain to the common ground shared between the addressees and the locutors and, therefore, they could not be computed by the sentiment analysis systems at hand.

	<b>Original tweet</b>	<b>Translation into English</b>	<b>nrc polarity</b>	<b>TextBlob polarity</b>
(9)	I tre gretini di Ultima Generazione sono stati già rilasciati	The three cretins with "g" of Last Generation have already been absolved	-0.07	0.0
(10)	Le bimbettole arcobalenate della merda buttata addosso agli altri [...]	The little rainbow girls of shit thrown at others [...]	-0.66	-0.1
(11)	Questi non sono ecoterroristi come al senato?	Aren't these eco-terrorists as in the senate?	-0.4	-0.5
(12)	[...] gli ecovandali imbrattano il Senato	[...] the ecovandals soil the Senate	-0.33	-0.4

For example, *gretini* (see section 2) and *bimbette arcobalenate* ‘little rainbow girls’ imply unpleasant and questionable stereotypes which may sound as follows: activists are childish people, not very smart, and live in an ideal world without connections to reality. Moreover, these stereotypes represent (9) and (10) as hilarious or facetious tweets [25] that increases the locutor’s pleasantness. Generally, the more pleasant the locutor, the greater their persuasive force. Instead, *ecoterroristi* ‘eco-terrorists’ (11) and *ecovandali* ‘eco-vandals’ (12) implicitly identify the actions performed by environmental activists as detrimental, and activists as members of a dangerous social group. However, this implicit interpretative level was ignored by the computerized sentiment analysis tools we used (and also by ChatGPT) that, therefore, were not always able to account for the complex rhetorical-argumentative strategies and stereotypes underlying the use of slurs.

## 5 Conclusion

In this paper, we compared different NLP tools at performing sentiment analysis on a corpus of Italian tweets related to the environmental blitz that occurred at the Senate in Rome on January 2<sup>nd</sup>, 2023. Our mixed-method approach, combining quantitative and qualitative analyses, indicated that while lexicon-based and Naïve Bayes classifier techniques provided interesting insights regarding the episode considered, they fell short in addressing inherent and notorious challenges for sentiment analysis, such as implicit language and indirect rhetorical-argumentative strategies. More sophisticated language models, such as ChatGPT, tackled these issues efficiently, by deciphering irony and sarcasm. In general, the results of our analyses indicated that the users’ attitudes were directed toward a variety of subjects and entities (the act itself, Italian politics, institutions, etc.) The sentiment expressed towards these entities varied within the corpus, without a clear majority position or stance in the Twitter debate. Nonetheless, upon conducting an in-depth qualitative manual analysis of a sample of 49 tweets, it appeared that the act of staining the Senate walls generated primarily negative sentiment among users.

## References

- [1] Jelani Ince et al. The social media response to Black Lives Matter: how Twitter users interact with Black Lives Matter through hashtag use. // *Ethnic and Racial Studies*. – 2017. – Vol. 40, no. 11. – P. 1814-1830.
- [2] Ali Rao Hamza et al. A large-scale sentiment analysis of tweets pertaining to the 2020 US presidential election // *Journal of Big Data*. – 2022. – Vol. 9, article 79. – P. 1-12.
- [3] Kruspe Anna et al. Cross-language sentiment analysis of European Twitter messages during the COVID-19 pandemic. // *Proceedings of the 1st Workshop on NLP for COVID-19 at the 58th Annual Meeting of the Association for Computational Linguistics*. – 2020. – P. 1-10.
- [4] Barrie Christopher and Ho Justin Chun-ting. *academicwitterR*: An R package to access the Twitter Academic Research Product Track v2 API endpoint. // *Journal of Open Source Software*. – 2021. – Vol. 6, no. 62, article 3272. – P. 1-2.
- [5] R Core Team. *R: A language and environment for statistical computing*. – 2022. – Vienna: R Foundation for Statistical Computing.
- [6] Sanguinetti Manuela et al. PoSTWITA-UD: An Italian Twitter Treebank in Universal Dependencies. // *Proceedings of the Eleventh International Conference on Language Resources and Evaluation – LREC2018*. – 2018. – P. 1397-1404.
- [7] Sanguinetti Manuela et al. Treebanking user-generated content: a UD based overview of guidelines, corpora and unified recommendations. // *Proceedings of the Twelfth International Conference on Language Resources and Evaluation – LREC2020*. – 2020. – P. 5240-5250.
- [8] Wijffels Jan. *udpipe*: Tokenization, parts of speech tagging, lemmatization and dependency parsing with the UDPipe NLP Toolkit (Version 0.8.11). – 2023. – CRAN.
- [9] Wickham Hadley. *Tidy Data*. // *Journal of Statistical Software*. 2014. – Vol. 59, no. 10. – P. 1-23.
- [10] Mohammad Saif M. and Turney Peter D. Crowdsourcing a Word-Emotion Association Lexicon. // *Computational Intelligence*. – 2013. – Vol. 29. – P. 436-465.
- [11] Jockers Matthew. *Syuzhet*: Extracts Sentiment and Sentiment-Derived Plot Arcs from Text (Version 1.0.6.). – 2020. – CRAN.
- [12] Silge Julia and Robinson David. *tidytext*: Text Mining and Analysis Using Tidy Data Principles. // *Journal of Open Source Software*. – 2016. – Vol. 1, no. 3. – P. 1-3.
- [13] Plutchik Robert. *The Emotions*. – 1991. – Lanham: University Press of America.

- [14] Plutchik Robert. The Nature of Emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. // *American Scientist*. – 2001. – Vol. 4. – P. 344–350.
- [15] Loria Steven. TextBlob: Simplified Text Processing. (Version 0.16.0). – 2020. – GitHub.
- [16] Bird Steven et al. Natural Language Processing with Python. – 2009. – Sebastopol: O’Reilly Media.
- [17] Han SuHun. Googletrans: Free and Unlimited Python Library for Google Translate API. (Version 3.0.0). – 2020. – GitHub.
- [18] Sazzed Salim. Cross-lingual sentiment classification in low-resource Bengali language. // *Proceedings of the Sixth Workshop on Noisy User-generated Text - W-NUT 2020*. – 2020. – P. 50-60.
- [19] Hunter John D. Matplotlib: A 2D graphics environment. // *Computing in Science & Engineering*. – 2007. – Vol. 9, no. 3. – P. 90-95.
- [20] Dmitry Davidov et al. Semi-supervised recognition of sarcastic sentences in twitter and amazon // *Proceedings of the fourteenth conference on computational natural language learning* – 2011. – P. 107-116.
- [21] Reyes Antonio et al. From humor recognition to irony detection: The figurative language of social media. // *Data & Knowledge Engineering* – 2012. – vol. 74. – P. 1–12.
- [22] Sbisà Marina. Detto non detto. – 2007. – Roma-Bari: Laterza.
- [23] Grice Herbert Paul. *Logica e conversazione. Saggi su intenzione, significato e comunicazione*. – 1993. – Bologna: Il Mulino.
- [24] Colston Herbert and Gibbs Raymond. A brief history of irony // Gibbs Raymond and Colston Herbert (eds.). *Irony in Language and Thought*. – 2007. – London: Routledge. – P. 3-24.
- [25] Gianti Andrea et al. Annotating irony in a novel Italian corpus for sentiment analysis // *Proceedings of the Workshop on Corpora for Research on Emotion Sentiment and Social Signals – LREC 2012*. – 2012. – P. 1-7.
- [26] Shelley Cameron. The bicoherence theory of situational irony // Gibbs Raymond and Colston Herbert (eds.). *Irony in Language and Thought*. – 2007. – London: Routledge. – P. 531-578.
- [27] Lucariello Joan. Situational irony: A concept of events gone away // Gibbs Raymond and Colston Herbert (eds.). *Irony in Language and Thought*. – 2007. – London: Routledge. – P. 467-498.
- [28] Dan Sperber and Deirdre Wilson. *Relevance: communication and cognition*. – 1986. – Cambridge: Harvard University Press.
- [29] Jeshion Robin. Slurs and stereotypes // *Analytic Philosophy*. – 2007. – vol. 54, no. 3. – P. 314-329.
- [30] Lombardi Vallauri Edoardo. *La lingua disonesta. Contenuti impliciti e strategie di persuasione*. – 2019. – Bologna: Il Mulino.