Stance Analysis of Twitter Users: the Case of the Vaccination Topic in Italy

This is the accepted manuscript of the paper with DOI: https://doi.org/10.1109/MIS.2020.3044968
This version has been made publicly available only after the embargo period requested by the journal.

Alessio Bechini[*‡], **Pietro Ducange**[*], **Francesco Marcelloni**[*], **and Alessandro Renda**[* †]
[*] Department of Information Engineering, University of Pisa, Largo Lucio Lazzarino 1, 56122 Pisa, Italy
[†] Department of Information Engineering, University of Firenze, Via di S. Marta 3, 50139, Firenze Italy
[‡] Corresponding author alessio.bechini@unipi.it

***Abstract*—People's opinion around social and political issues is currently witnessed by messages ordinarily posted on social media. With reference to a specific case study, namely the vaccination topic in Italy, this work discusses a crucial aspect in structuring the data processing pipeline in intelligent systems aimed at monitoring the public opinion through Twitter messages: a plain analysis of tweet contents is not sufficient to grasp the diversity of behavior across users. To get a sharper picture of the public opinion as expressed on social media, user-related information must be incorporated in the analysis. Relying on a dataset of tweets about vaccination and on an established text classification system, we present the results of a stance monitoring campaign with advanced analysis on temporal and spatial scales. The overall methodological workflow provides a sound solution for public opinion assessment from Twitter data.**

Index Terms: Automatic Stance Detection, Social Media Analysis, Twitter Users Analysis, Monitoring Attitude Towards Vaccination

■ **INTRODUCTION** Nowadays, Social Media produce an overwhelming amount of data, which represents a valuable source of information about the users who generate it and about society in general. For example, on the Twitter platform every day millions of users worldwide post short messages (popularly known as *tweets*) to share thoughts, viewpoints, and opinions about specific topics or events in their everyday life. Despite the noisy and irregular textual nature of tweets, they can be successfully analyzed with proper *text mining* techniques, exploiting methods of data mining, machine learning, and NLP (Natural Language Processing). Among the wide range of application fields, the monitoring of political and social concerns is currently attracting more and more interest [1], [2], mostly because information mined from the web represents a crucial semantic asset, along with more traditional opinion polls [3], for the estimation of the average stance of the population towards a given issue.

In the last years the vaccination topic has stirred up the social and political debate in Italy: although solid scientific evidence exists on vaccination benefits, other controversial opinions have gained popularity, drawing attention to possible negative side-effects of vaccines, or about putative pressures by pharmaceutical companies to increase their profits. The relevance of the topic has prompted us to investigate novel AI approaches able to provide a reliable view of the trend of public opinion from Twitter data:

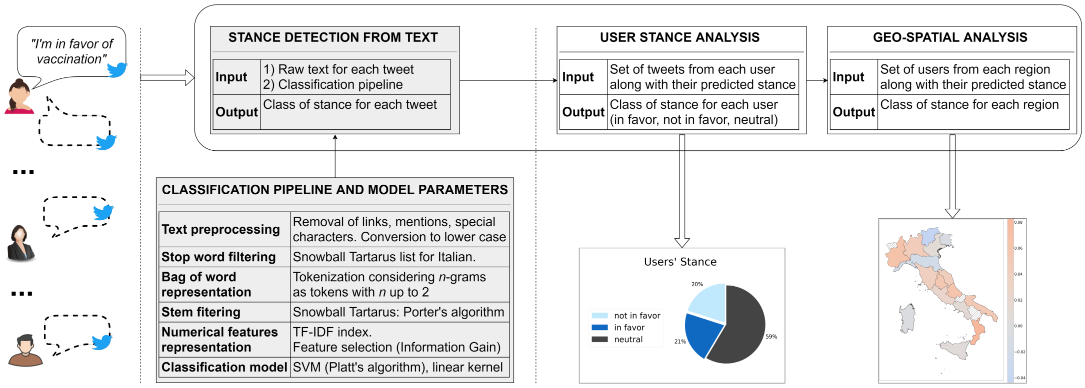


Figure 1. The methodological workflow for accurate stance analysis on social media: gray blocks, i.e. the *stance detection from text* module and the related *classification pipeline and model parameters*, have been previously described in [4]. The other downstream blocks, i.e *user stance analysis* and *geo-spatial analysis*, support the fine-grained analysis described in the text.

the availability of an intelligent system that supports such analysis may play an important role for Public Healthcare Organizations to promote awareness campaigns. In fact, it has been shown that opinionated messages posted on Twitter may actually impact the public opinion [5], and thus individual decision-making as well, with possible implications on public health.

Our work is pivoted on an automatic stance detection system that exploits NLP and AI techniques for tweets classification. We have discussed the constructive detail of the classification tool in a recent pilot study [4]. For the purpose of public opinion sensing, however, the automatic analysis cannot solely focus on the polarity of single tweets, but should move towards detecting the stance of *individual* users. Statistical considerations on users' activity and the introduction of ad-hoc modules in the methodological workflow depicted in Figure 1, elevate the investigation from the raw text-level semantics to a human- and society-level semantics [6]. With reference to the case study of the vaccination topic, we perform a retrospective analysis that shows how essential it is including user-related information in the stance monitoring system. As a result, a more realistic estimate of the public opinion is obtained along two dimensions: a temporal analysis provides indications about possible opinion shifts, while a spatial analysis carried out on a regional level captures its local variability.

THE VACCINATION DATASET AND STANCE CLASSIFICATION FROM TWEETS

The presented work stems from the availability of data obtained via an intelligent system that automatically infers the stance of Twitter contents about the vaccination topic in Italy [4]. The possibility to mine more accurate and detailed information out of collected data has driven the subsequent efforts.

The Vaccination Dataset

We carried out a monitoring campaign over a period of very intense public discussions on the vaccination topic, from September 1st, 2016 till June 30th, 2017. We specified a list of 38 Italian vaccine-related keywords, such as *#iovaccino* (hashtag for "I vaccinate") or *rischio vaccinale* (vaccine risk); then, we filtered out non-Italian tweets and duplicates. At the end, this setup allowed us to collect 99803 tweet objects, i.e. text documents with relative metadata.

The Stance Classification System from Text and the Monitoring Campaign

The stance detection task can be traced back to a three-class classification problem [7]: for each tweet the goal is to assess the position expressed towards the appropriateness of vaccination, namely *in favor*, *not in favor*, or *neutral*.

The design of our stance classification pipeline, including text preprocessing, numerical representation and model parameters, has been thoroughly described in one [4] of our previous works and is summarized in the dedicated block in Fig. 1. Details on the used algorithms are available also through references in the same paper. The final scheme adopts a bag-of-words with TF-IDF index (term frequency - inverse document frequency) representation for the tweet text, and a Support Vector Machine with linear kernel as classification algorithm; it is the result of an extensive model selection phase based on 10-fold stratified cross-validation on a dataset with 693 labelled tweets from the first three months of the monitoring campaign (neutral: 219, not in favor: 219, in favor: 255). The selected model outperformed a variety of methods, spanning from classical machine learning approaches to cutting-edge deep learning techniques, achieving an accuracy of 64.84%, which is in line with the performance reported in the literature for analogous case studies [8].

The use of the selected classification model during the monitoring period in the subsequent months lets us assess the daily distribution of tweets over the three classes. In general, the prevalent class is the neutral one, and the average stance shows a day-by-day variation, in accordance with social context-related events happened in Italy [4]; among them, we can mention:

- November 22nd, 2016: Approval of a law on vaccination requirements for school children in the Emilia Romagna Region, Italy;
- January 26th, 2017: Agreement between Italian Health Minister and Italian Regions about vaccinations requirement;
- March 16th, 2017: News on 230% increase of measles cases in Italy;
- April 17th, 2017: The Italian TV broadcasts a report on the vaccines controversy;
- April 19th, 2017: Fake vaccinations in the city of Treviso;
- June 22nd, 2017: Kid sick of leukemia died for measles in Monza.

TWITTER VACCINATION DATASET: A STATISTICAL OVERVIEW

According to the observations about the importance of explicitly considering users in our analysis, our dataset has to be described also from this viewpoint. First of all, we must underline that we are interested in *active* users, i.e. those who have posted at least one tweet on the specific day. The daily number of both tweets and active users along the whole observation period is reported in Figure 2. Unsurprisingly, the number of active users on a daily basis is coherent with the number of posted tweets. The average daily number of tweets per user over the whole time window is 1.18 ± 0.11.

The most evident characteristic in Figure 2 is the presence of several peaks, which can be traced back to events of public interest that took place in Italy [4]. The relative dates are reported on the x-axis of the chart.

A crucial step for an accurate, unsolicited, public opinion poll on Twitter consists in identifying users *actually relevant* for our analysis: we aim to tell apart the general public, mainly composed by individuals, from Twitter users corresponding to news-wires, news agencies or public authorities. We supposed that this distinction could be successfully accomplished by checking the *verified* attribute of the user object, as provided by the Twitter API: it indicates whether a user has been verified by the platform or not, i.e. if it is deemed to be an account of public interest.

In our dataset, out of a total of 28560 distinct active users, only 513 (1.8%) are *verified* ones. By checking usernames in this group, we noticed that most of them, around 300, cannot be referred to real people, but rather to news pages. The different nature of *verified* and *unverified* accounts may explain the gap between the two groups in terms of average number of tweets posted during the whole time span (verified: 8.2, unverified: 3.4). Such a result is further supported by the analysis of the distribution of the number of tweets per user (Figure 3). A visual analysis of the frequencies, reported on a logarithmic scale, highlights how the majority of the users have posted just one or a few tweets about the topic. However, the distribution tails suggest that the number of users who produced a high number of tweets are proportionally more frequent in the *verified* group than in the other. This observation can be quantitatively substantiated by fitting the histograms with power-law distributions of the form $pdf(x) \sim x^{-\alpha}, \alpha > 0$: the verified group has an *heavier* tail distribution, according to its lower value of α.

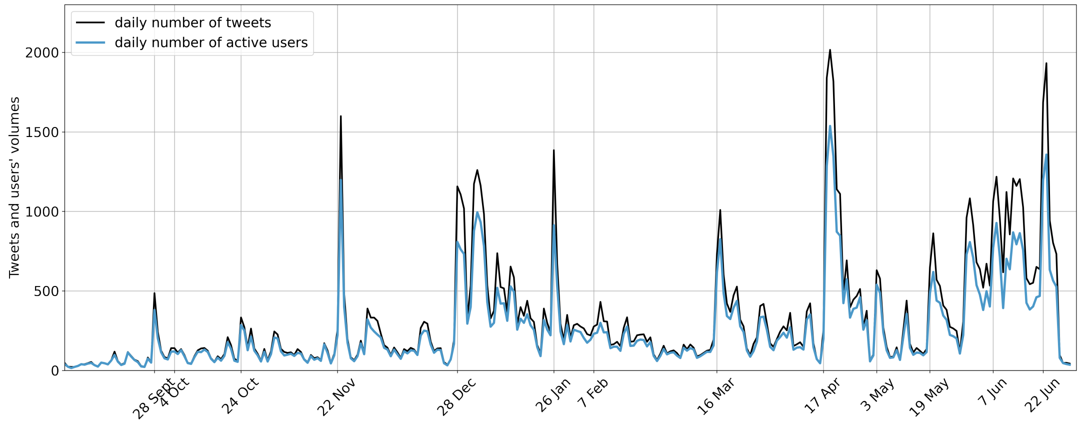


Figure 2. Daily number of tweets (upper line) and active users (lower line) over the whole observation period. Dates on the x-axis indicate occurrences of influencing social context-related events.

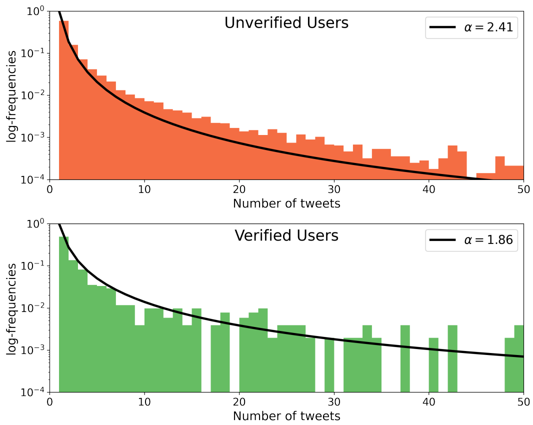


Figure 3. Normalized histograms of the users activity (number of tweets posted per day) for the *verified* and *unverified* groups. The solid black lines represent the best fitting power-law distributions, with the relative parameter α: the lower the value of α, the *heavier* the distribution tail. The logarithmic scale for the y-axis emphasizes the different shapes of the distribution tails for the two groups.

Finally, on the basis of the previous observations, we can reasonably assume that the unverified group might better reflect the actual opinions expressed by people. Of course, the analysis outcomes could be perceivably affected by the exclusion of verified users only if their contribution in the overall dataset is significant.

USER-CENTRIC STANCE ANALYSIS

In view of the above outcomes and with the aim of better characterizing the polarity of the public opinion over time, we devise a further analysis: our purpose is to capture the stance of individual users, rather than a rough picture of the stance of the overall volume of tweets posted about the vaccination topic.

Definition of User Stance

According to agreed positions in the research community [7], the stance of a user can be qualitatively defined as the publicly expressed position towards a proposition or a target. In our case study, the target of interest is pre-chosen and it consists in the appropriateness of vaccination. Since any active user can post multiple tweets per day, possibly with high variability over time, we have to recur to a stance characterization based on the stance of any single user in the dataset, referring to the contents he or she published on the topic. We adopt a recently proposed definition of the stance score [1]: let *Positive*, *Negative* and *Total* be the number of positive tweets (in favor of vaccination), negative tweets (not in favor of vaccination), and all tweets produced by a user in the time span of our monitoring campaign, respectively. Then, the stance score of a user u can be simply formulated as

$$\text{UserScore}_u = \frac{\text{Positive}_u - \text{Negative}_u}{\text{Total}_u} \tag{1}$$

The *UserScore* ranges between -1 and 1, and is negative for users not in favor of vaccination and positive for users in favor.

On the basis of the stance score of a user, we identify *UserStance*, a discretization of the *UserScore* in three classes, namely *Neutral (Neut)*, *InFavor (Pos)* and *NotInFavor (Neg)*. *UserStance* formulation enables a more interpretable representation of the position of the user, but it requires the definition of an arbitrary threshold σ. *UserStance* is defined as follows:

$$\text{UserStance}_u = \begin{cases} \text{Pos} & \textit{if} \quad \text{UserScore}_u > \sigma \\ \text{Neut} & \textit{if} \quad |\text{UserScore}_u| \leq \sigma \\ \text{Neg} & \textit{if} \quad \text{UserScore}_u < -\sigma \end{cases} \tag{2}$$

The choice of the exact value for σ is somehow arbitrary, and it cannot ignore psychological considerations on the perception of stance by human observers. In our experiments, driven by such consideration, we set $\sigma = 0.2$. We deemed it appropriate choosing a low threshold value, thereby capturing the polarity of a user even at intermediate values of stance score; however, the value cannot be too close to 0, because the neutral class should contain users who did not show a clear position throughout all their posted tweets. The final choice stems also from statistical considerations: it is reasonable assuming that some *UserScore* values occur more frequently than others (e.g. ± 0.5, ± 0.33, ± 0.25, ± 0.2), because most of the users posted a moderate number of tweets in the whole time window. Therefore, looking at all the possible combinations of 3, 4 and 5 tweets, we chose σ as the highest value obtained for a combination judged as neutral.

User Stance Analysis

As a first step of our user-oriented approach, we carry out the tweet stance analysis during specific events, taking also care of telling apart posts produced by the two different groups of users. Figure 4a shows how tweets from verified users (V) are in general less subjective (positively or negatively opinionated) than those from unverified users (NV). Furthermore, although the average number of tweets from verified users is higher, their inclusion does not significantly affect the analysis results obtained previously [4]: distributions obtained by considering all users (A) are in line with those obtained by limiting the survey to unverified users only.

In order to get a sharper picture of the public opinion, we base our analysis on *users' stance* instead of *tweets stance*, adopting the definition in Eq. 2. Results are reported in Figure 4b.

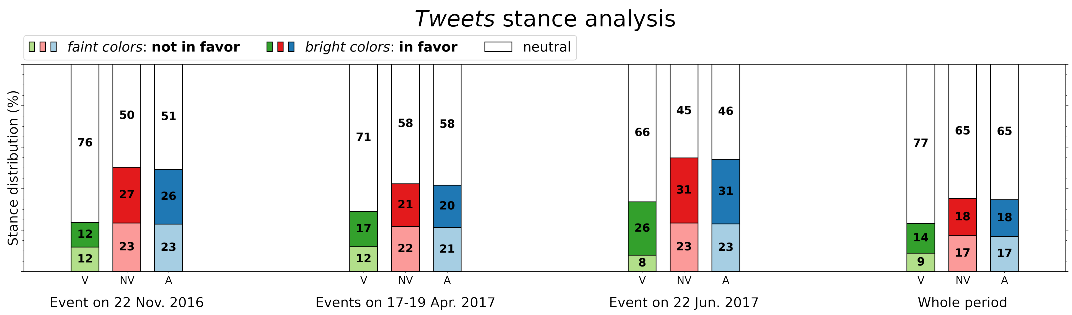
As expected, a comparison between Figures 4a and 4b highlights that measuring the public opinion solely on the basis of tweets is not entirely informative of the number of opinionated users. Such an outcome is more evident over the whole 10-months observation window, since it is more likely that a user has tweeted several times about the topic. In particular, 21% of unverified users are *in favor* of vaccination, while the overall percentage of positive tweets is 18%. On the other hand, negatively opinionated unverified users are 20% of the total, versus the 17% of the volume of negative tweets. Similar considerations stand for verified users: 27% of users are opinionated (18% positively, 9% negatively), against the 23% of subjective tweets posted by the same group (14% positive, 9% negative). In general, the lower number of *verified* than *unverified* users, who show either a positive or a negative opinion, further supports the validity of the assumption that the *verified* flag distinguishes public interest entities from individual users: in fact, in the annotation guidelines we have assigned the neutral label to generic news tweets.

For a comprehensive view of the public opinion over time, we evaluated the cumulative number of users in the three stance classes: for each day, the user stance is assessed from the tweets she/he has posted until that day. Figure 5a shows the obtained values for the unverified group. The trend of each class can be directly derived from the stacked histogram. In Figure 5b we report the trend of the polarity of public opinion (*P.O.*), simply evaluated as:

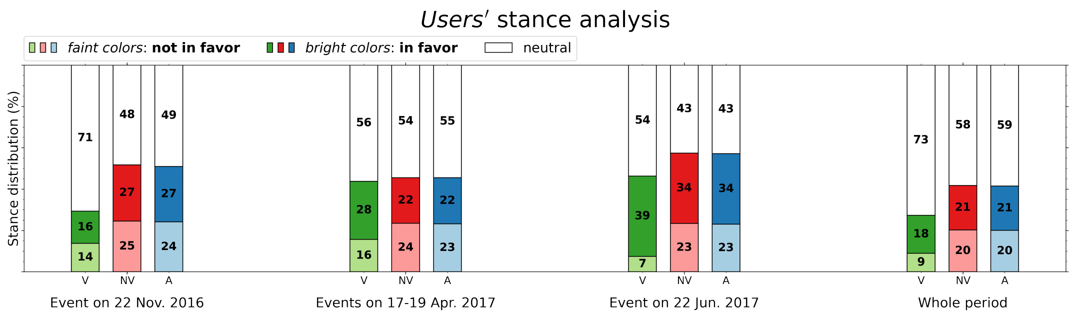
$$\text{P.O.}(T) = \sum_{t=0}^{T} \text{InFavor}(t) - \sum_{t=0}^{T} \text{NotInFavor}(t) \tag{3}$$

where "InFavor" and "NotInFavor" stands for the number of users in favor and against vaccination, respectively, and T varies along the whole monitoring period.

As an example, it is possible to notice that on



(a)



(b)

Figure 4. Distribution of the stance over the three classes during some relevant events and throughout the whole monitoring campaign. Distributions of *tweets* (4a) and *users* (4b) are reported. Results are shown for both the verified (V) and the unverified (NV) groups, and for all the users (A) as well.

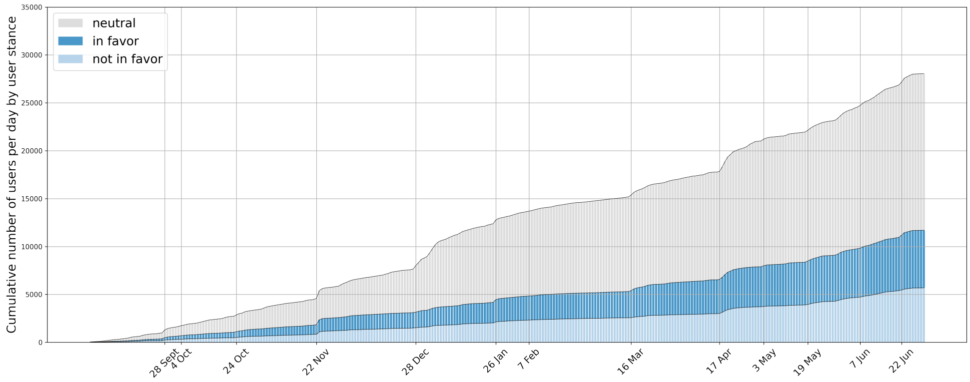
March 16th, 2017, the number of users in favor of vaccination grows *faster* than that of the opposite class. A similar pattern can be found on June 22nd, whereas the opposite trend occurs on April 17th, when the curve shows a slightly negative slope. We argue that the news about the increase of the measles cases (230%) (March 16th) and the death of a kid for measles (June 22nd) raised the awareness of the importance of vaccines, thus producing the fast increase of favorable users. On the other hand, on April 17th an episode of an Italian investigative journalism TV show was devoted to the controversies around vaccines: the hypothesis of a conflict of interest behind mandatory vaccines likely pushed many Italians to express a negative opinion about the topic. The subsequent decrease can instead be related to the events on May 19th and June 7th, both associated with a decree about the vaccination requirement in Italian schools.

Geospatial Analysis

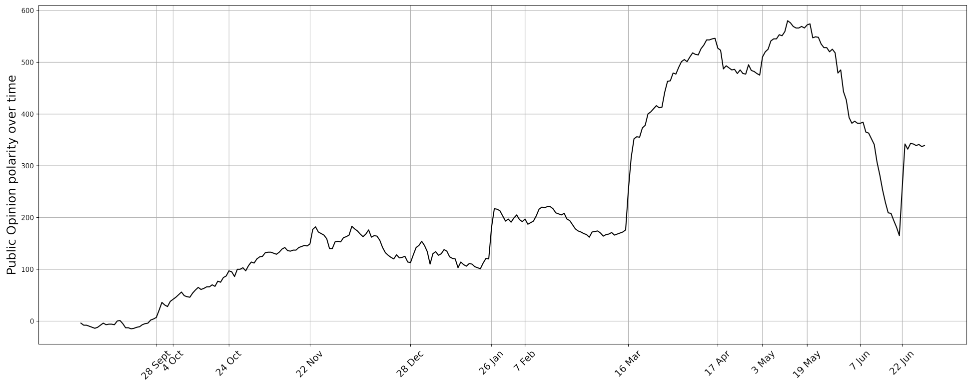
A nation-wide measure of the overall population stance could hide discrepancies that appear at a regional scale: lowering the analysis granularity to a regional or macro-regional level may highlight possible local nuances [9]. Hence, as a further analysis step we have considered the aspect of the *localization* of active users in our target case study.

Finding the exact spatial information of each tweet is not a trivial task, because only a very small number of users enables geolocatization of tweets in his/her preference settings. For this reason, we exploit the information, optional and approximate, occasionally provided by the user during the sign-in procedure: in this case the attribute *location* of the user object is not null, but the value may neither necessarily represent a real position, nor be machine-readable.

We found out that 65404 out of the 99803 tweet objects have a non-null *location* attribute. Overall, the total number of distinct *location*



(a) Stacked histogram showing the cumulative number of unverified users involved in the vaccination discussion over time, grouped by stance class.



(b) Public Opinion polarity over time, evaluated as the difference of the number of *in favor* users (blue bars in the plot at the top) and the number of *not in favor* users (light blue bars).

Figure 5. Trend of public opinion over time.

entries is 8237. For our spatial analysis, it is necessary to parse the string provided by the user and translate it into spatial coordinates (latitude and longitude). The GeoPy Python library fulfills this purpose, with the implementation of procedures for resolving a location from a string.

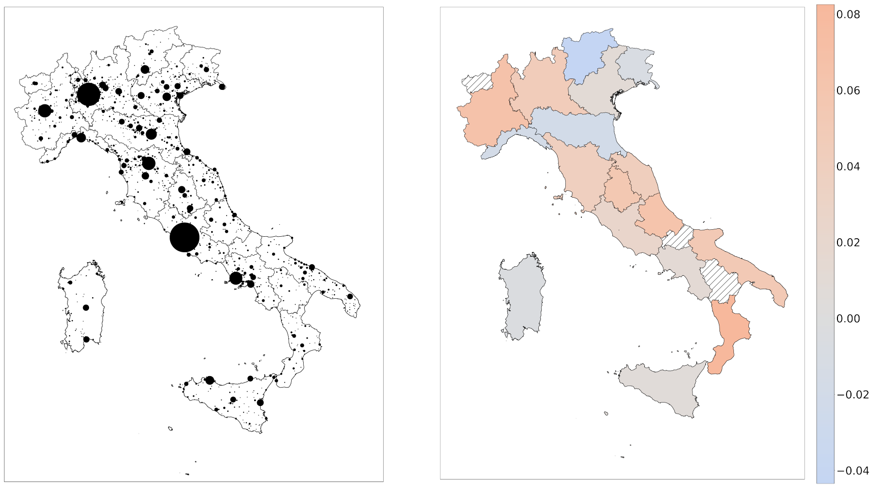
After filtering out invalid locations, we ended up with 42538 tweets (out of 99803), sent by 11527 distinct users (out of 28560), located in 2127 distinct positions. Figure 6a shows the distribution of active users on the Italian map: the bigger the marker size, the higher the number of users posting from that location. Map suggests that the metropolitan areas of Rome and Milan represent the most prolific spots. Moreover, the highest density of users occurs in central and northern Italy.

We have compared the distribution of the stance of users with and without a declared location. Table 1 shows the aggregated values for the group of verified users (V), unverified users (NV), and all users (All).

Table 1. Stance discrepancies between users with a known location (Loc) and without a known location (Not)

		Nr. of Users	Negative	Positive	Neutral
V	Loc	204	13.24	21.57	65.20
	Not	309	6.15	16.18	77.67
NV	Loc	11207	19.17	21.79	59.04
	Not	16840	20.99	21.25	57.76
All	Loc	11411	19.06	21.79	59.15
	Not	17149	20.72	21.16	58.12

Table 1 suggests that the sample of *verified* users who declare their location is not representative of the entire group. Indeed, in Figure 6b, an aggregated measure of the stance towards the vaccination topic over Italian regions is reported



(a) Geographic distribution of active users (based on self-declared location) on the Italian map

(b) Stance score on the administrative regional basis obtained according to Eq. 4, considering *unverified* users. Hatched pattern indicates regions with a low number of samples (less than 100 users).

Figure 6. Geospatial Analysis

for the *unverified* users. The stance score of a region r (*RegionalScore*) is evaluated in the same way as the users' stance score and by considering data from the whole monitoring campaign:

$$\text{RegionalScore}_r = \frac{\text{InFavor}_r - \text{NotInFavor}_r}{\text{Total}_r} \quad (4)$$

where the difference between the number of opinionated users (*InFavor* and *NotInFavor*) is normalized by the total number of users referring to the specific region.

The map reported in Figure 6b suggests that the prevailing stance slightly varies from region to region, but is moderately positive in most of the country.

RELATED WORKS

The vast majority of works that apply AI to stance detection do not take advantage of a "user-centric" approach. For example, this occurs in an interesting recent work [10] that discusses a new architecture for stance detection based on a memory network which separately memorizes text and contextual information, i.e. common sense knowledge. Conversely, another significant related study addresses the stance of Twitter *users* [1]: it deals with the offline measuring of stance and influence of Twitter accounts with regard to one single event (the Brexit referendum), rather than encompassing a fine grained temporal and geospatial characterization of public opinion.

Text analysis is not the only method that has been used to determine users' stance, and a recent work [11] proposes to make use of other basic linguistic (e.g. hashtags) and interaction (e.g. retweeted accounts) features. To this aim, other types of ML techniques have been exploited, i.e. dimensionality reduction and clustering.

CONCLUSION

An ever-growing number of studies exploit text mining to extract knowledge out of contents posted on Social Networks. Nevertheless, whenever the goal is to probe the public opinion on a certain topic, it should be considered that the text-level semantics does not necessarily reflect the multifaceted human-level and society-level semantics. This work makes use of Twitter data as information source, and introduces a novel methodological workflow for a fine-grained and more plausible characterization of the public opinion about the vaccination topic along time. As such, it represents an enabling technology for advances in the challenging field of context-sensitive stance detection.

We relied on a dataset collected in a period of very intense discussions on the target topic, and on a specific text classification system.

In the present work, we have found that the behaviour in terms of stance and number of posted tweets is different for accounts belonging to the group of *verified* users, which are imputable to accounts of public interest (newswires, public authorities), and for those indicated as *unverified* users. Significantly, by focusing on the temporal dimension of the collected data and the relative inferred user-centric stance, it is possible getting detailed insights into public opinion shifts, possibly associated with context-related events. Moreover, by resorting to the noisy information occasionally provided by users about their location, we have been able to perform a geospatial analysis of the matter: thematic maps of Italian country have been drawn, highlighting the location of active users and the aggregated stance at regional level.

From a structural viewpoint, all the user-centric analysis steps require the support of dedicated software modules, as part of the complete monitoring and classification system. The described methodological workflow provides the guidelines for their integration in even more effective intelligent systems for fine-grained, smart and low-cost analysis of public opinion from Twitter data.

With specific regard to the vaccination topic, this innovative type of monitoring tool may represent a powerful means for Public Healthcare Organizations to promote awareness campaigns and proper countermeasures against outbreaks of both current and eradicated diseases.

ACKNOWLEDGMENT

This work was partially supported by the Italian Ministry of Education and Research (MIUR) in the framework of the CrossLab project (Departments of Excellence).

REFERENCES

1. M. Grčar, D. Cherepnalkoski, I. Mozetič, and P. K. Novak, "Stance and influence of Twitter users regarding the Brexit referendum," *Computational Social Networks*, vol. 4, no. 6, 2017.
2. G. Bello-Orgaz, J. Hernandez-Castro, and D. Camacho, "Detecting discussion communities on vaccination in Twitter," *Future Generation Computer Systems*, vol. 66, pp. 125–136, 2017.
3. A. Tsakalidis, S. Papadopoulos, A. I. Cristea, and Y. Kompatsiaris, "Predicting elections for multiple countries using Twitter and polls," *IEEE Intelligent Systems*, vol. 30, no. 2, pp. 10–17, 2015.
4. E. D'Andrea, P. Ducange, A. Bechini, A. Renda, and F. Marcelloni, "Monitoring the public opinion about the vaccination topic from tweets analysis," *Expert Systems with Applications*, vol. 116, pp. 209–226, 2019.
5. M. Salathé, D. Q. Vu, S. Khandelwal, and D. R. Hunter, "The dynamics of health behavior sentiments on a large online social network," *EPJ Data Science*, vol. 2, no. 1, p. 4, 2013.
6. H. Zhuge, "Interactive semantics," *Artificial Intelligence*, vol. 174, no. 2, pp. 190–204, 2010.
7. D. Küçük and F. Can, "Stance detection: A survey," *ACM Comput. Surv.*, vol. 53, no. 1, Feb. 2020.
8. S. M. Mohammad, P. Sobhani, and S. Kiritchenko, "Stance and sentiment in tweets," *ACM Transactions on Internet Technology (TOIT)*, vol. 17, no. 3, p. 26, 2017.
9. G. Caldarelli, A. Chessa, F. Pammolli, G. Pompa, M. Puliga, M. Riccaboni, and G. Riotta, "A multi-level geographical study of Italian political elections from Twitter data," *PloS one*, vol. 9, no. 5, p. e95809, 2014.
10. J. Du, L. Gui, R. Xu, Y. Xia, and X. Wang, "Commonsense knowledge enhanced memory network for stance classification," *IEEE Intelligent Systems*, vol. 35, no. 4, pp. 102–109, 2020.
11. K. Darwish, P. Stefanov, M. Aupetit, and P. Nakov, "Unsupervised user stance detection on twitter," *Proc. of the Int'l AAAI Conference on Web and Social Media*, vol. 14, no. 1, pp. 141–152, May 2020.

Alessio Bechini is a researcher at the Department of Information Engineering of the University of Pisa, Italy. His research interests span the fields of concurrent and distributed systems, enterprise information systems, data management and integration, service management, with particular interest in bioinformatics problems. Recently, his work has been focused on issues in data mining for Big Data. He earned his Laurea degree in Electronics Engineering and the PhD degree in Information Engineering both from the University of Pisa in 1996 and 2003. He can be contacted at alessio.bechini@unipi.it

Pietro Ducange is an associate professor of Information Systems and Technologies at the Department of Information Engineering of the University of Pisa, Italy. His main research interests include evolutionary fuzzy systems, big data mining, social sensing and sentiment analysis. He received the M.Sc. degree in Computer Engineering and the Ph.D. degree in Information Engineering from the University of Pisa in 2005 and 2009, respectively. He can be contacted at pietro.ducange@unipi.it

Francesco Marcelloni is currently a full professor at the Department of Information Engineering of the University of Pisa, Italy. His main research interests include data mining for big data, sentiment analysis and opinion mining, multi-objective evolutionary algorithms, genetic fuzzy systems, fuzzy clustering algorithms, and data compression and aggregation in wireless sensor networks. He received the Laurea degree in Electronics Engineering and the Ph.D. degree in Computer Engineering from the University of Pisa in 1991 and 1996, respectively. He can be contacted at francesco.marcelloni@unipi.it

Alessandro Renda is currently a Ph.D. student at the Department of Information Engineering of the University of Pisa, Italy, enrolled in the Smart Computing programme, jointly awarded by the Universities of Pisa, Florence and Siena, Italy. His current research interests include machine learning algorithms for data streams, applications of deep learning methodologies and affective computing. He received the M.Sc. degree in Biomedical Engineering from the University of Pisa, Italy, in 2017. He can be contacted at alessandro.renda@unifi.it