

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

## Stance Evolution and Twitter Interactions in an Italian Political Debate

### **This is the author's manuscript**

*Original Citation:*

*Availability:*

This version is available <http://hdl.handle.net/2318/1669103> since 2018-05-30T10:49:10Z

*Publisher:*

Springer International Publishing AG, part of Springer Nature 2018

*Published version:*

DOI:10.1007/978-3-319-91947-8\_2

*Terms of use:*

Open Access


Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

# Stance Evolution and Twitter Interactions in an Italian Political Debate<sup>\*</sup>

Mirko Lai<sup>1,2</sup>, Viviana Patti<sup>1</sup>, Giancarlo Ruffo<sup>1</sup>, and Paolo Rosso<sup>2</sup>

<sup>1</sup> Dipartimento di Informatica, Università degli Studi di Torino

<sup>2</sup> PRHLT Research Center, Universitat Politècnica de València  
{lai,patti,ruffo}@di.unito.it, proso@dsic.upv.es

**Abstract.** The number of communications and messages generated by users on social media platforms has progressively increased in the last years. Therefore, the issue of developing automated systems for a deep analysis of users' generated contents and interactions is becoming increasingly relevant. In particular, when we focus on the domain of online political debates, interest for the automatic classification of users' stance towards a given entity, like a controversial topic or a politician, within a polarized debate is significantly growing. In this paper we propose a new model for stance detection in Twitter, where authors' messages are not considered in isolation, but in a diachronic perspective for shedding light on users' opinion shift dynamics along the temporal axis. Moreover, different types of social network community, based on retweet, quote, and reply relations were analyzed, in order to extract network-based features to be included in our stance detection model. The model has been trained and evaluated on a corpus of Italian tweets where users were discussing on a highly polarized debate in Italy, i.e. the 2016 referendum on the reform of the Italian Constitution. The development of a new annotated corpus for stance is described. Analysis and classification experiments show that network-based features help in detecting stance and confirm the importance of modeling stance in a diachronic perspective.

**Keywords:** Stance · Political Debates · Homophily · Twitter

## 1 Introduction

Nowadays, social media are gaining a very significant role in public debates. Political leaders use social media to directly communicate with the citizens, and citizens often take part in the political discussion, by supporting or criticizing their opinions or proposals. Therefore, social media provide a powerful experimental tool to deduce public opinion's mood and dynamics, to monitor political sentiment, and in particular to detect users' stance towards specific issues, like political elections or reforms, and their evolution during the debate and the

---

<sup>\*</sup> The work of the last author was partially funded by the Spanish MINECO under the research project SomEMBED (TIN2015-71147-C2-1-P).

related events [3]. Online debates are featured by specific characteristics. As observed by Adamic and Glance, web users tend to belong to social communities segregated along partisan lines [1]. Albeit the scientific debate is still open, some recent studies suggest that the so called “echo chambers” and “filter bubbles” effects tend to reinforce people’s pre-existing beliefs, and they also filter and censure divergent ones [13].

In this study we examine the political debate in Twitter about the Italian constitutional referendum held on December 4, 2016 in Italy. To carry on our analysis, we first collected a dataset of about 1M of Italian tweets posted by more than 100K users between November 24 and December 7, 2016, about the Italian constitutional referendum. Then, we extended the collection by retrieving retweets, quotes, and replies, aiming at a representation of political communication through different types of social networks. Furthermore, we divided our dataset in four temporal phases delimited by significant events occurred around the consultation period, for analyzing the dynamism of both users’ stance and social relations. We manually annotated the evolution of the users’ stance towards the referendum of 248 users, creating a corpus for stance detection (SD), i.e. the task of automatically determining whether the author of a text is in favour, against, or neutral towards a given target [9]. On this corpus, we were able to analyze the relations that occur among users not only considering the social network structure, but also the users’ stance. Based on this analysis we propose a new model for SD in Twitter featured by two main characteristics: (i) network-based features have been included in the model, which result from the analysis of different types of social network communities, based on retweet, quote, and reply relations; (ii) authors’ messages are not considered in isolation, but in a diachronic perspective. The major contributions of this work are:

1. *A new resource.* We developed a manually annotated corpus for SD about an Italian political debate, CONREF-STANCE-ITA henceforth. Such kind of resource is currently missing for Italian, in spite of the growing interest in the SD witnessed by the recent shared tasks proposed for English [9], Spanish and Catalan [12].
2. *Stance detection.* We propose a new SD model including a set of features based on social network knowledge. Experiments show that analyzing users’ relations helps in detecting stance.
3. *Stance diachronic evolution.* Our analysis on the debate provides some evidence that users reveal their stance in different ways depending on the stage of the debate; in particular, our impression is that users tend to be less explicit in expressing their stance as the outcome of the vote approaches.
4. *Network analysis.* Users tend to communicate with similar users, and a strong signal of homophily by stance among supporters and critics of the reform has emerged. Moreover, users having different opinions on the referendum often communicate using *replies*: a significant number of replies posted among ideologically opposed users occurs in the corpus.

The rest of the paper is organized as follows. In Section 2 we briefly discuss the related work. In Section 3 we describe the development of the corpus, and

its characteristics in terms of social network. In Section 4 we describe our SD model and the classification experiments. Section 5 concludes the paper.

## 2 Related Work

**Political sentiment and stance detection.** Techniques for sentiment analysis and opinion mining are often exploited to monitor people’s mood extracting information from users’ generated contents in social media [11]. However, especially when the analysis concerns the political domain [3], a recent trend is to focus on finer-grained tasks, such as SD, where the main aim is detecting users’ stance towards a particular target entity. The first shared task on SD in Twitter held at *SemEval 2016*, Task 6 [9], where is described as follows: “Given a tweet text and a target entity (person, organization, movement, policy, etc.), automatic natural language systems must determine whether the tweeter is *in favor* of the target, *against* the given target, or whether neither inference is likely”. Standard text classification features such as n-grams and word embedding vectors were exploited by the majority of the participants of the task. The best result was obtained by a deep learning approach based on a recurrent neural network [14].

Machine learning algorithms and deep learning approaches were also exploited in a second shared task held at *IberEval 2017* on gender and SD in tweets on Catalan Independence, with a focus on Spanish and Catalan [12]. With regard to SD, participating teams exploited different kinds of features such as bag of words, bag of parts-of-speech, n-grams, word length, number of words, number of hashtags, number of words starting with capital letters, and so on. The best result was obtained by a support vector machine (SVM) classifier exploiting three groups of features: *Stylistic* (bag of: n-grams, char-grams, part-of-speech labels, and lemmas), *Structural* (hashtags, mentions, uppercase characters, punctuation marks, and the length of the tweet), and *Contextual* (the language of each tweet and information coming from the URL in each tweet) [5].

**Political debates and diachronic perspective.** Recently, Lai et al. [6] explored stance towards BREXIT at user level by aggregating tweets posted by the same user on 24-hours time windows. This shows how stance may change after relevant events, a finding supported by the work of Messina et al. analysing the same debate [8]. A way to represent a dynamic system is aggregating empirical data over time considering different size of time-windows. Albeit the aggregation time window size is often dictated by the availability of data gathered and this issue has often been neglected in the literature, the importance of the choice of time-windows needs to be considered [4].

**Political debate and social media.** The huge amount of users generated data allows researchers to observe social phenomena with computational tools in an unprecedented way. Despite social media ease the access to a range of several conflicting views, some works suggest that the existence of the so called “echo chambers” (i.e., when users are exposed only to information from like-minded ones) and “filter bubbles” (i.e., when content is selected by algorithms according to the user’s previous behaviors) can have both positive and negative effects in

online and offline forms of political participation [1, 13]. Lazarsfeld and Merto theorized that homophily is involved [7] after the observation that people tend to bond in communities with others who think in similar ways, regardless of any differences in their status characteristics (i.e. gender, age, social status). Recent works shed some light on the relation between social media network structure and sentiment information extracted from posted contents. For example, Lai et al. [6] reported some preliminary results showing that a strong relation exists between user’s stance and friend-based social media community the user belongs, studying the English debate on BREXIT.

### 3 The ConRef-STANCE-ita Corpus

#### 3.1 Data Collection and Diachronic Perspective

Twitter is a microblogging platform where users post short messages called *tweets*. Users can share with their *followers* (users who follow them) the tweets written by other users; this type of shared tweets is known as *retweets*. Furthermore, users can add their own comments before retweeting making a tweet a *quote*. Moreover, it is possible to answer to another person’s tweet, generating a so called *reply*. Replying to other replies makes possible the development of longer *conversation threads*, including direct and nested replies.

Researches on Twitter are made easy by the Twitter’s REST and Streaming APIs, a set of clearly defined Web services that allow the communication between the Twitter platform and developers. All APIs return a message in JSON, a cross-platform data-interchange format. Also for these reasons, we chose Twitter as platform to gather our experimental data.

**Collection.** We collected tweets on topic of the Referendum held in Italy on December 4, 2016, about a reform of the Italian Constitution. On Sunday 4 December 2016, Italians were asked whether they approve a constitutional law that amends the Constitution to reform the composition and powers of the Parliament, the division of powers between the State, the regions, and other administrative entities. This referendum was source of high polarization in Italy and the outcome caused a sort of political earthquake<sup>3</sup>. The data collection consists of four steps:

1. About 900K tweets were collected between Nov. 24th and Dec. 7th through the Twitter’s Stream API, using as keywords the following hastags: #referendumcostituzionale, #iovotosi, #iovotono<sup>4</sup>.
2. The source tweet from each retweet was recovered by exploring the tweet embedded within the JSON field *retweeted\_status*. Then, we used the *statuses/retweets/:id* Twitter REST API in order to collect all retweets of the each retweeted tweet present in the dataset.

<sup>3</sup> The majority of the voters rejected the reform causing the resignation of Matteo Renzi, the Prime Minister that assumed full responsibility for the referendum defeat.

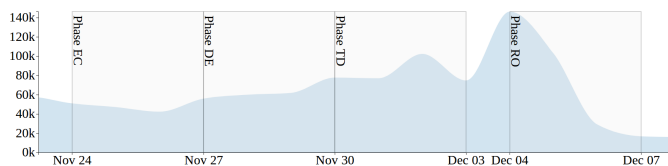
<sup>4</sup> #constitutionalreferendum, #Ivoteyes, #Ivoteno

3. We recovered the quoted tweet of each quote exploring the embedded tweet within the JSON field *quoted\_status*.
4. We retrieved *conversation threads* recursively resorting to the Twitter REST API *statuses/show/:id*, by using, as parameter, the *id* specified in the field *in\_reply\_to\_status\_id* of each replied tweet.

Through these steps, we have thus enlarged the available number of tweets (more than 2M) w.r.t. those gathered by the Twitter Stream API alone (about 900K). Therefore, we extended the number of possible relations between users (retweets, quotes, and replies) involved in the debate through steps 2, 3 and 4 for deeper analyzing social media networks.

**Diachronic Perspective.** Using the same methodology described in [6], we divided the collected tweets in four discrete temporal phases, each one delimited by significant daily spikes of tweets. The spikes correspond to events occurred leading up to the referendum, as it is shown in Figure 1. We thus consider the following four 72-hour temporal phases:

- “The Economist” (EC): The newspaper *The Economist* sided with the “yes” campaign of the referendum (tweets retrieved between 2016-11-24 00:00 and 2016-11-26 23:59).
- “Demonstration” (DE): A demonstration supporting the “no” campaign of the referendum had been held in Rome exactly one week before the referendum (tweets retrieved between 2016-11-27 00:00 and 2016-11-29 23:59).
- “TV debates” (TD): The Italian Prime Minister, Matteo Renzi, who supported the “yes” campaign of the referendum, participated to two influential debates on TV (tweets retrieved between 2016-11-30 00:00 and 2016-12-02 23:59).
- “Referendum outcome” (RO): The phase includes the formalization of the referendum outcome, and the resignation of the Italian Prime Minister (tweets between 2016-12-04 00:00 and 2016-12-06 23:59).



**Fig. 1.** Daily frequency of tweets and the discrete division in temporal phases.

### 3.2 Annotation for Stance

We applied to our data the same annotation schema previously exploited at the shared tasks proposed at SemEval 2016 [9] and IberEval 2017 [12] for annotating stance in English, Spanish and Catalan tweets. Here three labels were considered: FAVOR, AGAINST, NONE. The annotation guidelines provided to the annotators follow.

From reading the following tweets, which of the options below is most likely to be true about the tweeter’s stance or outlook towards the reform subjected to the Italian Constitutional referendum?

- **FAVOR**: We can infer from the tweet that the tweeter supports the reform.
- **AGAINST**: We can infer from the tweet that the tweeter is against the reform.
- **NONE**: We can infer from the tweet that the tweeter has a neutral stance towards reform or there is no clue in the tweet to reveal the stance of the tweeter towards the reform.

**Stance at user level.** We followed the same approach described in [6], where the stance is at user level rather than at tweet level. This means that we deduced the stance from multiple texts written by the same user rather than considering the stance of a single text. We define a *triplet* as a set of three tweets written by the same user in a single temporal phase. The triplet includes: one tweet, one retweet and one reply. This means that each user, for which we annotated the stance, may be a connected node in a network of relations of both retweet or reply. The users who wrote at least one tweet, one retweet, and one reply (a triplet) in each temporal phase are 248. The annotated corpus consists of 992 triplets (248 users by 4 temporal phases). For example, a single user wrote the tweet, the retweet, and the reply highlighted by the black bullet. The reply message in the triplet includes also the related tweet (marked with a white bullet) written by another user.

TWEET	<ul style="list-style-type: none"> <li>• Travaglio: “Il 2 dicembre grande serata nostra Costituzione in diretta streaming” #ioDicoNo URL via @fattoquotidiano (Travaglio: “The 2th December a great night for our Constitution in streaming live” #ISayNo URL through @fattoquotidiano)</li> </ul>
RETWEET	<ul style="list-style-type: none"> <li>• RT @ComitatoDelNO: Brava @GiorgiaMeloni che ricorda a @matteoreenzi di (provare a) dire la verità almeno 1 volta su 10! (RT @NOCCommittee: well done @GiorgiaMeloni who reminds to @matteoreenzi to (try to) say the truth at least 1 time over 10!)</li> </ul>
REPLY	<ul style="list-style-type: none"> <li>• @angelinascanu @AntonellaGramig @Rainbowit66 per la poltrona. La cosa più cara a voi del #bastaunSi #IoDicoNo #IoVotoNO #vergognaPD (@angelinascanu @AntonellaGramig @Rainbowit66 for their seats. The most important thing for you of the #justaYES #ISayNo #IVoteNO #shamePD)</li> </ul>
↪ TO	<ul style="list-style-type: none"> <li>oGià dovrebbe spiegare...ma la risposta si conosce. Il 4 dicembre #bastaunSi #IoVotoSI URL (He already should explain... but the answer is known. The 4 December #justaYES #IVoteYES URL)</li> </ul>

**Manual Annotation.** Two native Italian speakers, domain experts, provided two independent annotations on all the 992 triplets. For what concerns the

triplets for which an agreement between the first two annotators was not achieved, we recurred to CrowdFlower<sup>5</sup>, a crowd-sourcing platform. We exploited 100 tweets as test questions in order to evaluate the CrowdFlower annotators. We required that annotators were native Italian speakers living in Italy. The annotators have been evaluated over the test questions and only if their precision was above 80% they were included in the task. A further annotator was required unless at least 60% of the previous annotators agreed on the stance of a given triplet. We required a maximum of 3 additional annotators in addition to the 2 domain experts, regarding ambiguous triples. Overall, each triplet was annotated by at least 2 annotators to a maximum of 5.

**Agreement.** We calculated the inter-annotation agreement (*IAA*) as the number of annotators who agree over the majority label divided by the total number of annotators for each single triplet. This type of inter-annotator agreement was proposed by Mohammad et al. [10] to overcome the problem of calculating agreement over a set of documents annotated by a different number of annotators. The *IAA* calculated over all 992 triplets is 74.6%. Finally, we discharged triplets annotated by 5 annotators having less than 3 annotators in agreement on the same label. We named the Twitter with the stance about the Constitutional reform as CONREF-STANCE-ITA, and it consists of 963 triplets.

**Label distribution.** Table 1 shows the label distribution over temporal phases in the CONREF-STANCE-ITA<sup>6</sup>. The percentage of triplets labeled as AGAINST is higher than the rest of labels. This is in tune with the final outcome of the referendum (59.12% vote “no”)<sup>7</sup>. The frequency of the label NONE over the different temporal phases is another interesting point. As we can see, the distribution of this label constantly increases from phase EC to phase RO.

**Table 1.** Label distribution

LABEL	EC	DE	TD	RO	OVERALL
AGAINST	72.7%	72.7%	71.5%	62.8%	69.9%
FAVOR	19.8%	18.3%	16.9%	14.0%	17.2%
NONE	<b>6.2%</b>	<b>9.1%</b>	<b>11.6%</b>	<b>22.3%</b>	12.3%
disagreement	1.2%	0%	0%	0.8%	0.5%

We also explored if users’ stance changes over time. We find that 66.8% of the users were labeled with the same stance in all three intervals (55.0% AGAINST, 10.9% FAVOR, 0.8% NONE). For what concerns users that change stance across different time intervals, about 12% of them varies annotated stance in the last phase (10% AGAINST → NONE; 2.5% FAVOR → NONE). Similar

<sup>5</sup> <http://www.crowdfunder.com>

<sup>6</sup> ConRef-STANCE-ita and code available at: [github.com/mirkolai/ Stance-Evolution-and-Twitter-Interactions](https://github.com/mirkolai/Stance-Evolution-and-Twitter-Interactions).

<sup>7</sup> [https://en.wikipedia.org/wiki/Italian\\_constitutional\\_referendum,\\_2016](https://en.wikipedia.org/wiki/Italian_constitutional_referendum,_2016)



observations were made in [6], while investigating English tweets on the UK European Union membership referendum debate (BREXIT).

### 3.3 Social Media Networks Communities

*Networks Science* has applications in many disciplines due to networks (or graphs) that are able to represent complex relations among involved actors. Those relations are usually called *edges* and the actors are *nodes*. A network is *weighted* when each edge is characterized with a numerical label that reflects the strength of the connection between two nodes. Therefore, the network is *unweighted* when there is no difference between edges, i.e., all weights are equals to one.

In this work, we represent the relations among Twitter users involved in the Constitutional Referendum debate in the form of graphs. We extracted social media network communities from each graph using the Louvain Modularity algorithm [2]. Then, we examined the structure of four types of communication networks focusing on the dynamism of interactions and the percentage of *uncross-stance* relations (edges between two users with the same stance) for each type of communication. Table 2 shows the dimensions of each graph in each temporal phase.

**Table 2.** Graphs’ dimension for each temporal phases.

	RETWEET		QUOTE		REPLY	
	<i>nodes</i>	<i>edges</i>	<i>nodes</i>	<i>edges</i>	<i>nodes</i>	<i>edges</i>
Overall	94,445	405,843	24,976	69,240	20,936	41,292
EC	25,793	83,134	6,907	13,574	6,236	8,651
DE	28,015	98,717	7,577	15,665	6,663	9,714
TD	33,860	127,593	9,599	22,479	8,801	14,046
RO	63,805	158,243	14,919	21,977	8,497	10,832

**Retweet.** First, we consider the retweet-based networks. We gathered the retweet list of 649,306 tweets. We created a directed graph for each temporal phase. In particular, an edge between two users exists if one user retweeted a text of the other user during a defined temporal phase. The Louvain Modularity algorithm find about 800 communities for each temporal phase (except for the phase RO where about 1100 communities exist). About 90% of users belong to less than 20 communities.

**Quote.** We also considered the quote-based networks. We created a directed graph for each temporal phase. An edge between two users exists if one user quotes the other within a defined temporal phase. The four quote-based networks contain about 500 distinct communities (except for phase RO where about 800 communities exist). 1% of the communities contains about 50% of users.

**Reply.** Finally, we considered the reply-based networks. We recursively gathered the replied tweets of 81,321 replies. The recovered replies are 103,559 at the end of the procedure. Then, we created a directed graph for each temporal phase. In particular, an edge between two users exists if one user replies the other during a defined temporal phase. The communities extracted from the reply-based

network are about 700 for each temporal phase (except for phase RO where about 1500 communities exist). There are many communities that contain very few users, indeed only the 2% of the communities contains more than 10 users.

### 3.4 Relations and Stance

Here, we analyze the relations that occur among users not only considering the network structure, but also the users’ stance. Table 3 shows the percentage of “uncross-stance” relations (edges between two users with the same stance) considering only users annotated with the labels AGAINST or FAVOR. We considered both unweighted and weighted graphs, where the strength of the connection is the number of interactions (retweet, quote, or reply) between two users within the same temporal phase. Following, we evaluated the percentage of “uncross-stance” relations for each of the four network types.

**Table 3.** The percentage of uncross-stance relations among users.

	RETWEET		QUOTE		REPLY	
	<i>unweighted</i>	<i>weighted</i>	<i>unweighted</i>	<i>weighted</i>	<i>unweighted</i>	<i>weighted</i>
Overall	98.6%	99.1%	<b>94.8%</b>	<b>97.6%</b>	<b>81.9%</b>	<b>77.3%</b>
EC	98.1%	98.9%	94.0%	96.9%	82.0%	71.9%
DE	99.7%	99.8%	96.1%	97.9%	83.2%	81.0%
TD	98.6%	99.4%	93.9%	97.7%	81.2%	78.9%
RO	<b>97.5%</b>	<b>97.6%</b>	96.3%	97.9%	80.9%	77.1%

**Retweet.** First, we analyzed the reply-based network. The considered 3,099 relations are respectively distributed on the four temporal phases as follows: 749, 885, 989, and 476. The column RETWEET in Table 3 shows the percentage of uncross-stance retweets in the retweet-based network. The users usually retweet only tweets belonging to users having the same stance (98.6% and 99.1% overall respectively for unweighted and weighted graphs). There are no significant differences between unweighted and weighted graphs. Notably, the percentage of uncross-stance relations slightly decreases in the phase RO.

**Quote.** Then, we considered networks based on quote relations. We performed the analysis over 717 relations (respectively 183, 179, 247, and 108 for each temporal phase). The column QUOTE in Table 3 shows the percentage of uncross-stance quotes over the temporal phases. There are no significant differences between temporal phases, but the percentage of uncross-stance relations varies between unweighted and weighted graphs (from 94.8% to 97.6% overall).

**Reply.** Finally, we analyzed the reply-based network. 662 relations are distributed over the four temporal phases as follows: 172, 173, 207, and 110. The column REPLY in Table 3 shows the percentage of uncross-stance in both unweighted and weighted for each temporal phase. There are no significant differences between temporal phases, but the percentage of uncross-stance replies significantly varies between unweighted and weighted graphs (in particularly from 81.9% to

77.3% overall). Moreover, here we find a signal that uncross-stance relations is not the whole story.

## 4 Experiments

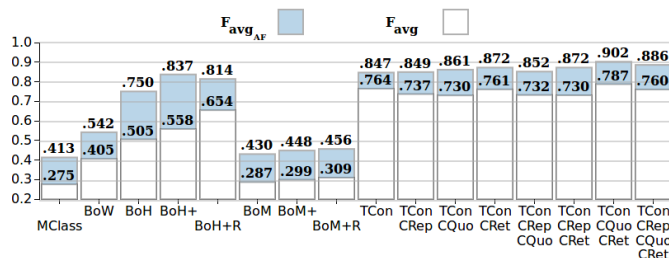
We propose a new SD model relying on a set of new features, which exploits SVM as machine learning algorithm in a supervised framework. As evaluation metrics, we use two macro-average of the  $F_{micro}$  metrics i.e.  $F_{avg}$  and  $F_{avg_{AF}}$ . The first one computes the average among f-AGAINST, f-FAVOR, and F-NONE  $F_{micro}$  metrics. The second one, proposed in both SemEval 2016 Task 6 and IberEval 2017 SD tasks [9, 12], computes the average between f-AGAINST and f-FAVOR  $F_{micro}$  metrics. We compare our results with two baselines such as: unigrams, bigrams and trigrams Bag of Words using SVM (*BoW*) and Majority Class (*MClass*). We compute the two metrics performing a five-cross validation on the CONREF-STANCE-ITA corpus employing each combination of the following features:

- **Bag of Hashtags** (*BoH*) and **Bag of Mentions** (*BoM*): hashtags/mentions as terms for building a vector with binary representation. These features use the texts contained in the tweet, the retweet, and the reply belonging to the triplet.
- **Bag of Hashtags+** (*BoH+*) and **Bag of Mention+** (*BoM+*): tokens extracted from the hashtags/mentions as terms for building a vector with binary representation. We segmented hashtags in tokens using the *greedy algorithm* attempting to find the longest word from a list of about 10M words extracted from Wikipedia’s Italian pages. We consider as token the lemma of the verb *to vote* when an inflection of this verb is found. For what concerns mentions, tokens are the result of the *name* splitting, using space as separator. Names are extracted from the *User Object* field *name* of the mentioned user. The feature uses the texts contained in the tweet, the retweet, and the reply belonging to the triplet.
- **Bag of Hashtags+ Replies** (*BoH+R*) and **Bag of Mentions+ Replies** (*BoM+R*): These features are similar to *BoH+* and *BoM+*, but they use information from the conversation thread, by exploiting the text of the replied tweet belonging to the triplet. A different prefix has been used in order to differentiate these tokens from the ones belonging to *BoH+* and *BoM+* feature.

The combination of *BoH+*, *BoM+*, and *BoH+R* (afterwards *TCon*) achieved the highest results,  $F_{avg}$  0.76 and  $F_{avg_{AF}}$  0.85. Notably, removing *BoH+R* from *TCon*,  $F_{avg_{AF}}$  declines to 0.83 and  $F_{avg}$  declines to 0.69. The model probably is benefiting from the opposition of stance between reply and replied tweets.

**Network-based Features.** In order to study the impact of knowledge of the social network for each network’s type, we introduced three new features that

consider the community which the user belongs to: **Retweet Communities** ( $CRet$ ), **Quote Communities** ( $CQuo$ ), and **Reply Communities** ( $CRep$ ) respectively. In particular, considering the temporal phase  $tp \in \{EC, DE, TD, RO\}$ ,  $N$  binary variables exist, one for each of the  $N$  detected communities in the retweet-based, quote-based, or reply-based networks. The variable set to one corresponds to the community to which the users who wrote the triplet belongs in the given temporal phase  $tp$ . Fig. 2 shows the combination of the three network-based features with  $TCon$ . As we can see, the combination of  $TCon$ ,  $CRet$ , and  $CQuo$  achieved the highest value for both  $F_{avg}$  and  $F_{avg_{AF}}$  (0.79 and 0.90, respectively) by improving the results obtained using only the  $TCon$  features (0.76 and 0.85, respectively). Nevertheless, adding the  $CRep$  feature does not improve neither  $F_{avg}$  and  $F_{avg_{AF}}$ .



**Fig. 2.** F measured achieved adding network-based features to  $TCon$ .  $F_{avg_{AF}}$ : average between f-AGAINST and f-FAVOR  $F_{micro}$  metrics.  $F_{avg}$ : average among f-AGAINST, f-FAVOR, and F-NONE  $F_{micro}$  metrics.

## 5 Discussion and Conclusion

In this work we created a manually annotated Italian corpus for addressing SD from a diachronic perspective, which allows us to shed some light on users' opinion shift dynamics. We observed that in this debate, users tend to be less explicit on their stance as the outcome of the vote approaches. Analyzing the relations among users, we also observed that the retweet-based networks achieved the highest percentage of uncross-stance relations (percentage very close to 100%). This is a signal that Twitter's users retweet almost exclusively tweets they agreed on. Very high percentage of uncross-stance were achieved also by the quote-based networks. The variation between unweighted and weighted graphs could mean that users mainly quote users they agree on. Therefore, it is more likely to be in agreement when the number of quotes connecting two users increases. Interestingly, the opposite is happening on reply-based networks, where we can observe a higher percentage of communications between users with different stances. These observations led us to propose a new model for SD, which includes three new network-based features. The performed experiments show that adding  $CRep$  and  $CQuo$  features to content-based features considerably improve the accuracy of

SD. We are guessing that when homophily is observed, the user’s awareness of being a member of a community can ease user’ stance prediction. This does not happen in *CRep*: although the users mainly reply to other users with a similar opinion, we observe about 20% of cross-stance edges. This is a particularly interesting case where inverse homophily (or also heterophily) could be observed. It will be matter of future investigations.

## References

1. Adamic, L.A., Glance, N.: The political blogosphere and the 2004 u.s. election: Divided they blog. In: Proceedings of the 3rd International Workshop on Link Discovery. pp. 36–43. LinkKDD ’05, ACM, New York, NY, USA (2005)
2. Blondel, V.D., Guillaume, J.L., Lambiotte, R., Lefebvre, E.: Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment* **10**, 10008 (2008). <https://doi.org/10.1088/1742-5468/2008/10/P10008>
3. Bosco, C., Patti, V.: Social media analysis for monitoring political sentiment. In: Alhajj, R., Rokne, J. (eds.) *ESNAM*, pp. 1–13. Springer (2017)
4. Krings, G., Karsai, M., Bernhardsson, S., Blondel, V.D., Saramäki, J.: Effects of time window size and placement on the structure of an aggregated communication network. *EPJ Data Science* **1**(1), 4 (2012). <https://doi.org/10.1140/epjds4>
5. Lai, M., Cignarella, A.T., Hernández Fariás, D.I.: ITACOS at ibereval2017: Detecting stance in Catalan and Spanish tweets. In: Proceedings of IberEval 2017. vol. 1881, pp. 185–192. CEUR-WS (2017)
6. Lai, M., Tambuscio, M., Patti, V., Ruffo, G., Rosso, P.: Extracting graph topological information and users’ opinion. In: Proceeding of CLEF 2017, pp. 112–118. Springer, Dublin, Ireland (2017). [https://doi.org/10.1007/978-3-319-65813-1\\_10](https://doi.org/10.1007/978-3-319-65813-1_10)
7. Lazarsfeld, P.F., Merton, R.K.: Friendship as a social process: A substantive and methodological analysis. In: Berger, M., Abel, T., Page, C. (eds.) *Freedom and Control in Modern Society*, pp. 18–66. Van Nostrand, NY, USA (1954)
8. Messina, E., Fersini, E., Zammit-Lucia, J.: All atwitter about Brexit: Lessons for the election campaigns. <https://radix.org.uk/work/atwitter-brexit-lessons-election-campaigns> (2017), online; accessed 28 January 2018
9. Mohammad, S., Kiritchenko, S., Sobhani, P., Zhu, X., Cherry, C.: Semeval-2016 task 6: Detecting stance in tweets. In: Proceedings of SemEval-2016. pp. 31–41. ACL, San Diego, California (2016)
10. Mohammad, S.M., Sobhani, P., Kiritchenko, S.: Stance and sentiment in tweets. *ACM TOIT* **17**(3), 26:1–26:23 (2017). <https://doi.org/10.1145/3003433>
11. Pang, B., Lee, L.: Opinion mining and sentiment analysis. *Found. Trends Inf. Retr.* **2**(1-2), 1–135 (2008). <https://doi.org/10.1561/1500000011>
12. Taulé, M., Martí, M.A., Rangel, F.M., Rosso, P., Bosco, C., Patti, V., et al.: Overview of the task on stance and gender detection in tweets on Catalan independence at IberEval 2017. In: Proceedings of IberEval 2017. vol. 1881, pp. 157–177. CEUR-WS (2017)
13. Theocharis, Y., Lowe, W.: Does Facebook increase political participation? Evidence from a field experiment. *Information, Communication & Society* **19**(10), 1465–1486 (2016)
14. Zarrella, G., Marsh, A.: Mitre at semeval-2016 task 6: Transfer learning for stance detection. In: Proceedings of SemEval-2016. pp. 458–463. ACL, San Diego, CA, USA (2016)