Information Fusion 52 (2019) 344–356

Contents lists available at ScienceDirect

Information Fusion

journal homepage: www.elsevier.com/locate/inffus

Full Length Article

Social context in sentiment analysis: Formal definition, overview of current trends and framework for comparison

J. Fernando Sánchez-Rada*, Carlos A. Iglesias

Intelligent Systems Group, Universidad Politécnica de Madrid, Spain

ARTICLE INFO

Keywords:
Sentiment analysis
Social context
Social network analysis
Online social networks

ABSTRACT

Sentiment analysis in social media is harder than in other types of text due to limitations such as abbreviations, jargon, and references to existing content or concepts. Nevertheless, social media provides more information beyond text, such as linked media, user reactions, and relations between users. We refer to this information as social context. Recent works have successfully leveraged the fusion of text with social context for sentiment analysis tasks. However, these works are usually limited to specific aspects of social context, and there have not been any attempts to analyze and apply social context systematically. This work aims to bridge this gap by providing three main contributions: 1) a formal definition of social context; 2) a framework for classifying and comparing approaches that use social context; 3) a review of existing works based on the defined framework.

1. Introduction

Recent years have witnessed the rise of social media. Platforms such as Twitter or Facebook have become the de facto way to share thoughts and opinions with a wide audience [41]. Studies of Twitter usage show that about 19% of tweets contain a reference to a brand or product, 20% of which also show some expression of brand sentiment [39]. As a consequence, companies and researchers have grown interested in social media as a way to monitor public opinion. The sheer amount of social media content makes it impractical or impossible to manually process it. Hence, automatic sentiment analysis has grown very popular.

Sentiment analysis has been applied for many years in other types of opinionated content, such as online reviews or news articles. However, social media content poses several unique challenges to natural language processing in general, and to sentiment analysis in particular [64]. Some of these challenges are imposed by the very nature of social media platforms, such as limited length and relying on associated media. Other difficulties are caused by the characteristics of human interaction in these types of media. e.g., short attention span, need for immediacy, and use of specialized language. The result is a type of text that is short, full of jargon or abbreviations, ephemeral, and rife with references to contextual information.

There are different approaches to sentiment analysis in social media [3,14,71]. Most techniques are content-centric. They exploit specific linguistic characteristics of social media, just like previous research has done for other media (e.g., news articles) and domains (e.g., movie reviews). Some works try to overcome abbreviations and short texts in social media by finding external sources to link text to, such as news articles [32] or Wikipedia pages [29]. Other works leverage the specific language in these media by finding cues for sentiment (e.g., smileys and hashtags) [21]. When the textual content is also accompanied by multimedia, such as images or videos, the sentiment information in these media obtained with multimodal analysis [69] may also be exploited.

Nevertheless, these approaches fail to use the fact that information shared on social networks is not isolated. The meaning of a particular piece of content (e.g., a Tweet, a Facebook status or a blog post) may only be understood when its context is taken into consideration. This context includes visible information such as previous content that belongs to the same conversation, previous interactions between users, or people that interacted with the content (e.g., by liking it). It also includes seemingly unrelated social features. For instance, some demographic factors such as age and gender have been shown to correlate with sentiment and vocabulary [89], and they have been used to improve sentiment classification [37].

New sentiment analysis techniques are starting to incorporate the fusion of information from text and social context. Social context has also been introduced in other fields related to sentiment analysis, such as spam detection, where clues to identify spammers are usually hidden in multiple aspects of context, such as previous content, behavior, relationship, and interaction [15]. Unfortunately, the definition of social features, the methods employed to extract them, and how they are applied to sentiment analysis tasks vary greatly from work to work. These differences in notation and approaches are taxing, which makes comparing different works harder.

* Corresponding author.
E-mail addresses: jf.sanchez@upm.es (J.F. Sánchez-Rada), carlosangel.iglesias@upm.es (C.A. Iglesias).

https://doi.org/10.1016/j.inffus.2019.05.003
Received 11 December 2018; Received in revised form 8 May 2019; Accepted 13 May 2019
Available online 13 May 2019

Thus, further research is needed to delve more deeply into the notion of social context and the fusion of social context with traditional textual sentiment analysis. This work seeks to answer the following questions:

- Q1. What is social context?
- Q2. Can social context improve sentiment analysis?
- Q3. What elements of social context are more relevant for sentiment analysis purposes?

As a result, the contributions herein are threefold. First, this work proposes a formal and general definition of social context. Secondly, a framework to compare existing works in the field is proposed. In this framework, each work is described using a multi-level taxonomy that classifies each approach in terms of the proposed definition of social context, and other factors such as the machine learning techniques applied. Thirdly, the state of the art in sentiment analysis using social context is organized and compared using the defined framework. Moreover, the results reported by each work in the analysis have been aggregated and analyzed, to simplify the comparison of approaches.

The remaining of this paper is structured as follows. Section 2 presents an overview of the state of the art in sentiment analysis prior to social context, and an introduction to social network analysis; Section 3 introduces a formal definition of social context; Section 4 presents the framework for comparison of approaches to sentiment analysis using social context; Section 5 provides an overview of the state of the art, using the framework presented in the previous section; Lastly, Section 6 discusses the main conclusions drawn from this work and future lines of research.

2. Related work

This section is overview of relevant work in the fields of sentiment analysis and social network analysis. Each field is discussed in a separate section. The former discusses different approaches in sentiment analysis, including deep learning and ensemble techniques. The latter introduces Social Network Analysis (SNA), and it focuses on community detection due to its importance in several of the works reviewed.

2.1. Sentiment analysis

Although sentiment analysis has been an active research topic for decades, it has grown in popularity with the advent of online opinion-rich resources [64]. In turn, these resources have also added their own set of limitations and challenges.

Over the last two decades, numerous works have explored sentiment analysis in different applications and using different approaches. These approaches can be grouped into machine learning, lexicon based, and hybrid [71]. Of the three, machine learning techniques and hybrid approaches seem to be dominant [3,65,90], and lexicon techniques are typically incorporated into machine learning approaches to improve their results. Machine learning approaches apply a predictor (a classifier, or an estimator) on a set of features that represent the input. The set of predictors is not very different from those used in other areas. Instead, the complexity in these approaches lies in extracting complex features from the text, filtering only relevant features, and selecting a good predictor [78].

One of the most straightforward features is the Bag Of Words (BOW) model. In BOW, each document is represented by the multiset (bag) of its constituent words. Word order is disrupted, and syntactic structures are broken. As a result, a great deal of information from natural language is lost [94]. Therefore, various types of features have been exploited, such as higher order n-grams [63]. A more sophisticated feature is Part of Speech (POS) tagging [30]. In it, a syntactic analysis process is run, and each word is labeled (tagged) with its syntactic function (e.g., noun). Additionally, syntactic trees can be calculated. Using these trees, the words in the input can be rearranged to a more convenient position while still conveying the same meaning. Note how these two types of features only rely on lexical and syntactical information. For this reason, they are sometimes referred to as surface forms.

Surface forms can also be combined with other prior information, such as word sentiment polarity [11,28,44,54,57]. This prior knowledge usually takes the form of sentiment lexicons, i.e., dictionaries that associate words in a domain or language with a sentiment. Some lexicons also include non-words such as emoticons [36,40] and emoji [60]. These alternative forms of writing have been shown very useful, as they can dominate textual cues and form a good proxy for text polarity [36].

The use of lexicon-based techniques has many advantages [82], most of which stem from their combination with other methods. For instance, it is possible to generate lexicons that are domain dependent or that incorporate language-dependent characteristics. Lexicons and syntactic information can also be combined with linguistic context to shift valence [68]. On the other hand, there are several disadvantages to lexicon approaches. First, creating lexicons is an arduous task, as it needs to be consistent and reliable [82]. It also needs to account for valence variability across domains, contexts, and languages. These dependencies make it hard to maintain domain-independent lexicons. An alternative to retain independence while encoding domain, language, and context variability is through semantic representation of the lexical resources in the form of ontologies. An ontology can encode both lexical [52] and affective [81] nuances, both in the lexicons and in the automatic annotations [9]. This is especially useful for aspect-based sentiment analysis, as the differences between aspects can be incorporated into the ontology [91].

In recent years, new approaches based on deep learning have shown excellent performance in Sentiment Analysis [5,19]. In contrast with traditional techniques, deep learning techniques learn complex features from data with minimum human interaction. These algorithms do not need to be passed manually crafted features: they automatically learn new complex features. The downside is that the quality of the features heavily depends on the size of the training data set. Hence, they often require large amounts of data, which is not always available. They also raise other concerns such as interpretability [49,51] or its inability to adapt to deal with edge cases [51]. In the realm of Natural Language Processing (NLP), most of the focus is on learning fixed-length word vector representations using neural language models [42]. These representations, also known as word embeddings, can then be fed into a deep learning classifier, or used with more traditional methods. One of the most popular approaches in this area is word2vec [55]. The downside of these methods is that they require enormous amounts of training data. Luckily, several researchers have already applied these methods to large corpora such as Wikipedia and released the resulting embeddings.

Lastly, it is also possible to combine independent predictors to achieve a more accurate and reliable model than any of the predictors on their own. This approach is known as ensemble learning. Many ensemble methods have been previously used for sentiment analysis. Ensemble methods can be classified according to two main dimensions Rokach [73]: how predictions are combined (rule-based and meta-learning), and how the learning process is done (concurrent and sequential). A new application of ensemble methods is the combination of traditional classifiers based on feature selection and deep learning approaches [3].

2.2. Social network analysis and community detection

Social Network Analysis (SNA) is the investigation of social structures [62]. It provides techniques to characterize and study the connections between people, and their interactions. SNA is not limited to Online Social Network (OSN), but to any kind of social structure. Other examples of social network would be a network of citations in publications or a network of relatives. Through SNA techniques, it is possible to extract information from a social network that may be useful for sentiment analysis, such as chains of influence between users, groups of like-minded users, or metrics of user importance.

There are several ways in which SNA techniques can be exploited in sentiment analysis, but most of them fall under one of two categories: those that transform the network into metrics or features that can be used to inform a classifier; and those that limit the analysis to certain groups or partitions of the network.

A simple example of metrics provided by SNA could be user's follower in-degree (number of users that follow the user) and out-degree (number of users followed by the user), which could be used as features for each user [79]. However, these metrics are not very rich, as they only cover users directly connected to a user, and it does so in a very naive way: all connections are treated equally. Other more sophisticated metrics could be used instead of in/out-degree, such as centrality, a measure of the importance of a node within a network topology, or PageRank, an iterative algorithm that weights connections by the importance of the originating user. Several works have introduced alternative metrics for user and content influence in a network [33,59].

The second category of approaches is what is known either as network partition or as community detection, depending on whether the groupings may overlap. Intuitively, community detection aims to find subgroups within a larger group. This grouping can be used to inform a classifier, or to limit the analysis to relevant groups only. More precisely, community detection identifies groups of vertices that are more densely connected to each other than to the rest of the network [66]. The motivation is to reduce the network into smaller parts that still retain some of the features of the bigger network. These communities may be formed due to different factors, depending on the type of link used to connect users, and the technique used to detect the communities. Each definition has its own set of characteristics and shortcomings. For instance, if users are connected after messaging each other, community detection may reveal groups of users that communicate with each other often [22]. By using friendship relations, community detection may also provide the groups of contacts of a user [25].

The reader is referred to other publications [61,66] for further details of the different definitions of community and algorithms to detect them.

3. A Definition of social context

This section introduces a novel definition of social context and its components. The definition is focused on OSN aspects, and it is based on previous definitions and on the observed usage of social context features in the state of the art.

Since the inception of Twitter and its API in late 2006, several works have used social features to complement text [6]. This section aims to introduce a general definition of social context that both encompasses existing definitions and formalizes the loose or implicit definitions used in most works.

To the best of our knowledge, the first formal definition of social context was introduced by Lu et al. [50]. They defined the social context of a set of Reviews R as *the triple* $C(R) = \langle U, A, S\rangle$, *of the set of reviewers U, the authorship function A, and the social network relation S*. Although their work is focused on reviews, it identifies the three main entities of this social context: the content (*review*), the content producer (*the author*) and the user-relations (*the social network relations*). Later works have also referred to social context in different terms [58,93], but a formal definition is seldom provided. For instance, Ren and Wu [72] define both Social Context and Topical Context, based on the graph of relations and their adjacency matrix. Namely, Social Context is defined as $G_S = \{u, S\}$, where u is the set of users and S is the adjacency matrix between users, and Topical Context is defined as $G_t = \{t, T\}$, where t is the set of topics, and T is the adjacency matrix of topics.

Based on these definitions, and our analysis of the state of the art, we have identified four types of elements that make up Social Context (Fig. 1): content (C), users (U), relations (R) and interactions (I). These elements are related as follows.

Users are connected through relations and interactions. Relations are stable connections between two or more users (R^u). There are multiple types of relations, such as friendship, or belonging to the same group. Some types of relations are undirected or mutual, like kinship, whereas others are directed or asymmetrical, such as liking and following relations. Interactions appear when a user communicates with others (I^u). The types of interactions include direct messages, replies, and user mentions. Most of these types also involve the creation of content. When a user creates or posts new content, an authorship relation between the user and the content is formed (R^{uc}). New content may also be related to existing content (e.g., as a reply or a mention, R^c), or to other users (e.g., the user is mentioned in the content, R^{uc}). Users may then interact with the newly created content (I^{uc}), by replying to it, liking it, saving it, etc.

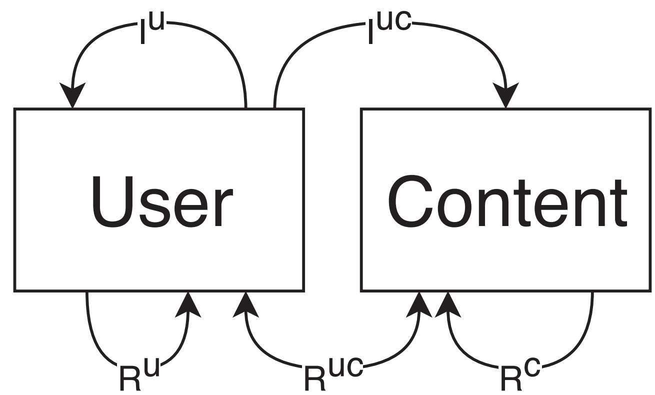

Fig. 1. Model of Social Context, including: content (C), users (U), relations (R^c, R^u and R^{uc}), and interactions (I^u and I^{uc}).

All elements are rich entities with different attributes. The specific attributes that can be used depend on the type of element and the OSN. Content attributes (e.g., text, creation date) and user attributes (e.g., name, age, gender) are commonly used. Although interaction and relation attributes are not as widespread, they are also important. They provide information such as when the interaction happened, or the weight of the relation. These attributes make it possible to filter specific connections, and to apply algorithms that rely on weighted graphs.

An additional concept to take into account is temporal dependence. New content is continuously created, and existing content is changed or removed. Relations are similar, as they are forged and dissolved naturally; and users can join, delete their accounts or become inactive. The relevance of social context variation over time is illustrated in Section 4.3 with the introduction of dynamic approaches.

These ideas about the elements of Social Context and their dynamic nature are condensed in the following definitions. First, Definition 1 covers Social Context as a whole and establishes its constituent elements.

Definition 1. Social Context is the collection of users, content, relations, and interactions which describe the environment in which social activity takes place. Namely:

$$SocialContext(\tau) = \langle C, U, R, I\rangle(\tau) = \langle C(\tau), U(\tau), R(\tau), I(\tau)\rangle$$

At any point in time τ: $C(\tau)$ is the set of content (Definition 2) generated by these users; $U(\tau)$ is the set of users (Definition 3); $I(\tau)$ is the set of interactions (Definition 5) between users, and of users with content; $R(\tau)$ is the set of relations (Definition 4) between users, between pieces of content, and between users and content.

This is a very general definition which only sets up the main elements, and it relies on the definition of each element to fully characterize context. To simplify the notation in the remaining definitions, time dependence will be implicit from here on: $SocialContext = \langle C, U, R, I\rangle$. This can be done without loss of generality. Whenever time dependence is relevant, we will refer to time-dependent social context as dynamic social context and to time-independent social context as static social context.

To illustrate the definitions, we will model an example of social context for a sentiment analysis task on Facebook content. For this analysis, we only need access to status updates by some users, and photos uploaded to a set of Facebook pages (groups).

The first element in social context is content:

Definition 2. The collection of content is defined as:

$$C = \{c_{t,i} \mid t \in T_c\} \tag{1}$$

Where T_c are all the types of content available, and each $c_{t,i}$ is a piece of content of a certain type t. Each piece of content should be unambiguously identified by its type and an identifier (i).

Our example context only includes two types of contents: status updates and photos. Each type of content may be given some attributes. Some of these attributes are common, such as the creation date. Others are specific for that type, such as the keywords for status updates, and the link to the image file for photos. Additionally, each photo and each status has to be given an identifier, which may also be the one given by the Facebook API. So far, the context defined is not very useful, as it would only allow us to analyze the sentiment of the status updates and the photos (using other modalities).

The next element in Social Context is the collection of users in the network.

Definition 3. Let the set of users be:

$$U = \{u_1, u_2, \dots, u_n\} \tag{2}$$

Where each u_i is a specific user that is unambiguously identified by its user identifier i. Each user may have one or more roles. The set of roles for a user is:

$$\rho(u_i) = \{t \mid \rho_t(u_i) = 1, u_i \in U, t \in T_\rho\} \tag{3}$$

Where T_ρ are all possible roles in a context, and $\rho_t(u_i)$ is a function that determines whether user u_i has been assigned role t.

Roles define the function of users within the network. They usually restrict the type of interactions and relations a user may have, and with what content and users. e.g., online fora have the role of topic moderators, in addition to regular users. The aim of moderators is to decide what content should be allowed, to edit it, and to manage users that misbehave. Hence, new relations (e.g., edited-by) and interactions (e.g., ban) are available to this specific role. If the user is a moderator of more than one topic, several roles will apply.

Our example context will include the profiles of the users in our study and their attributes. Since we are only interested in age and location, users will just have those attributes. Our users may also have roles. In our case, we will be interested in page administrators. At this point, the lack of connection between users and content hampers other types of analysis.

The categorization of connections in Social Context is based on the concept of social ties in the social sciences, i.e., dyadic relations [8]. Social ties are grouped into one of four categories: similarities, such as co-location or being the same gender; social relations, such as kinship (e.g., family ties), role (e.g., friendship), or affection (e.g., liking); interactions, such as having talked to each other, or harming one another; and flows, such as sharing information, beliefs, or resources. For the sake of simplicity, and based on the use of context in the state of the art, only two types of connections are modeled as part of Social Context: relations (Definition 4) and interactions (Definition 5). The remaining social ties (similarities and flows) can be modeled as an equivalent relation or interaction, depending on the case. Similarities are not typically considered as ties in themselves but rather as conditions or states that increase the probability of forming other kinds of ties. Flows are typically inferred from interactional and relational data [8] so, for the sake of simplicity, they can be thought of as another type of relation or interaction.

Hence, relations are connections such as friendship, kinship, group membership or liking each other, whereas interactions are connections such as getting in touch, re-sharing each other's content, etc. There are two main differences between relations and interactions that motivate their distinction. First, relations are few and slow-changing, whereas interactions are plentiful and short-lived. Secondly, content can be related to other content (e.g., a reply and the original content), while interactions are always performed by a user agent.

Formally, relations and interactions are defined as follows:

Definition 4. Given a set of content C, and a set of users U. Relations are the connections between users (R^u), between users and content (R^{uc}) and between different content (R^c). Formally:

$$R \equiv \{r_t \mid t \in T_r\} = R^u \cup R^{uc} \cup R^c \tag{4}$$

$$R^u_t = \{r^u_{t,u_i,u_j} \mid u_i, u_j \in U, u_i \neq u_j, t \in T_{r,u}\} \tag{5}$$

$$R^{uc}_t = \{r^{uc}_{t,u_i,c_j} \mid u_i \in U, c_j \in C, t \in T_{r,uc}\} \tag{6}$$

$$R^c_t = \{r^c_{t,c_i,c_j} \mid c_i, c_j \in C, c_i \neq c_j, t \in T_{r,c}\} \tag{7}$$

Where $T_{r,c}$ are the types of relations between two pieces of content, $T_{r,uc}$ are the types of relations between users and content, and $T_{r,u}$ are the types of relations between users.

Definition 5. Given a set of content C, and a set of users U. Interactions are the activities carried on by a user that involve either another user (I^u), or a piece of content (I^{uc}). Formally:

$$I \equiv \{i_t \mid t \in Ti\} = I^u \cup I^{uc} \tag{8}$$

$$I^u_t = \{i^u_{t,u_i,u_j,i} \mid u_i, u_j \in U, t \in T_{i,u}\} \tag{9}$$

$$I^{uc}_t = \{i^{uc}_{t,u_i,u_j,i} \mid u_i \in U, c_j \in C, t \in T_{i,uc}\} \tag{10}$$

Where $T_{i,uc}$ are the types of interactions between user and content, $T_{i,u}$ are the types of interactions between users, and i is an identifier for the interactions, as multiple interactions of the same type are possible.

With all elements defined, we can go back to the previous example of Social Context on Facebook. From the possible types of relations between users (R^u), we may add two: user friendship and kinship. These two relations would allow us to group users that are closely related. To link users with content, we will choose two types of user-content relations (R^{uc}): authorship, and mentions (i.e., the link between the content and the users it mentions). As for relations between content (R^c), we may choose replies (i.e., the link between two pieces of content when one mentions the other). Lastly, we will only have access to interactions between users and content (I^{uc}) in the form of likes, reactions, and replies. Due to technical limitations, we will not have access to user interactions, such as direct messages.

The resulting example context would allow for richer analyses that exploit information such as inferred groups of people based on how often they interact with each other or appear in photos together. Sentiment analysis may exploit prior knowledge about the sentiment of the user (via the authorship relation), or even knowledge about the sentiment of friends and acquaintances (through either relations or interactions between users). It may even be possible to find people within the group that have changed the opinion of the people with whom they interact.

Table 1 shows other types of user, content, relations and interactions found in popular OSN. It includes common elements in the OSN analyzed in the state of the art: Twitter, Weibo, Reddit, Facebook, blogging platforms and Wikipedia.

Table 1
Types of Social Context elements in different OSN.

OSN	Content (T_c)	User roles (T_ρ)	Relations (T_r)			Interactions (T_i)	
			User-User ($T_{r,u}$)	User-Content ($T_{r,uc}$)	Content-Content ($t_{r,c}$)	User-User ($t_{i,u}$)	User-Content ($t_{i,uc}$)
Twitter	Tweet	User	Follow Friend	Author Mentioned Favorite	Reply Retweet	Mention Reply	Reply Retweet Mention
Weibo	Weibo	User	Follow Friend	Author Mentioned Favorite	Reply Reshare	Mention Reply	Reply Reshare
Reddit	Post Comment	User Admin	Follow	Author Mentioned	Link Reply	Mention Reply	Vote Gild Reply Mention
Facebook	Status Page Comment Photo Event	User Page admin	Friend Relative	Author Admin Fan Own Tagged Attend Like React	Link Reply Contain	Mention Reply Tag	Comment Re-share
Blog	Post Comment	Author Reader	Follow	Author Like	Link Reply	Mention Reply	Reshare Comment
Wiki	Page Comment	Editor Reviewer	-	Author Edit Review	Link Parent Reply	-	Edit

The tabular format does not capture how different types of relations or interactions are unique to certain types of content and/or user roles. We will exemplify this fact using Facebook since it has different types of content and users roles. In Facebook, we may consider four main types of content. There are statuses, which are posts by users which are shown on their own profile (i.e., user feed). Statuses are very rich, they may mention other users, include location information, link to other content, or even express the mood of the author. The visibility of the status is governed by the user's privacy settings, and the relationship of the user to others. For instance, privacy-minded users may make their statuses only available to their close friends, while other users may make theirs public. Similarly, users can create pages, which are public profiles created around a specific topic, such as a business, a brand, or a cause. Pages are similar to user profiles, but they can be administered by one or more users. Another type of content is photos, which may be linked to a user profile or to a page. Photos can include information about the users that appear in them, which creates a relation between the photo and the users. Events are a different type of content that is used to organize gatherings and to give information about them. Users may indicate whether they will attend, comment on the event, and invite other users to join.

Users may interact with content to which they have access in different ways: by liking it; by commenting to it, which creates new content that other users may interact with; or by expressing their reaction or emotion to it, such as surprise. These types of interaction are common for all types of content. Some types of content provide other means of interaction, such as re-sharing of posts, which allows users to share a post by other user in their own profiles.

The primary means for interaction between users is through content, either by interacting with the content, e.g., users may reply to each other's content, by including other users in their content, e.g., by adding a mention in a comment or a tag in a photo. Lastly, they may interact through special actions such as poking each other, or through private instant messages. Since these interactions are private, they have not been included in the table.

Some researchers are concerned that the typical follower-friend relation might not be enough to capture the richness of relations in online media [20]. They also propose researching into new multifaceted approaches which take into consideration more aspects of the network simultaneously. Social context has been intentionally defined with those approaches in mind. The definition of Social Context can be interpreted in the form of sets, or in its equivalent graph form, where users and content are vertices, and both relations and interactions are edges. The graph form can be combined with different types of links (T_c, T_u, T_r, T_i) to generate multiplex networks [27] (i.e. a multilayered network of users and content), which can be exploited in multifaceted approaches.

To conclude, the usage of the social network [43] and the effect of the social network on user behaviour [18] depend on other aspects such as cultural differences, factual information and events. This type of information falls outside the scope of social context, and will need to be encoded through other means such as a knowledge graph, or a description of events. However, social context will capture information such as language of a user or creation time of content, which can be used to link the user or content to that external information. This concept will be further explained in Section 4.2.

4. Framework for research on social context in sentiment analysis

This section defines a novel framework to compare sentiment analysis approaches that exploit social context. The framework is centered around a multi-levelled taxonomy for structuring research in the field. The first level refers to the dataset used. The second level covers the scope of Social Context built from the dataset. The third level covers machine learning methods applied. The fourth level covers the type of social context used (static and dynamic). Each level is further explained in a separate section.

4.1. Dataset

The datasets used for analyzing social context can be identified by several characteristics. The first of them is the online social network from which the data was gathered. Twitter predominates in this area, due to its relatively open API and abundance of content. The second characteristic is the type of annotation on content. Likewise, the third characteristic is the type of annotation on users. In this work, we focus on sentiment (polarity), but other annotations such as stance, emotion, and quality of the content are often used. In the case of polarity, the classes used may also differ. i.e. positive (+), negative (-) and neutral (0). The fourth, fifth, and sixth characteristics are the type of link between users, between pieces of content, and between users and content. These links can stem either from a relation or from an interaction, as mentioned in the definition of social context.

4.2. Context scope

Researchers have to choose what information from their datasets to select for the social context in their work. They may also complement the original data with information from external sources. As a consequence, every work employs a different context. Nonetheless, a closer inspection reveals some patterns: some elements are commonly used together (e.g., users and friendships), and some elements are harder to obtain or rarer than others (e.g., follower-followee relations are more common than retweets or favorites). As contexts get more and more complex, they start including more unusual elements in addition to the more basic ones.

Hence, we propose a classification of works based on the complexity or scope of their context. Our proposal is inspired by the micro, meso and macro levels of analysis typically used in social sciences [7]. The

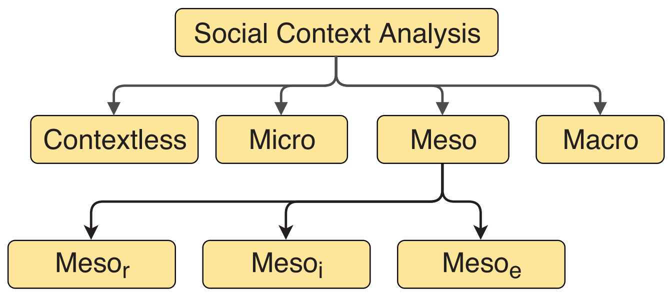

Fig. 2. Taxonomy of approaches, and the elements of Social Context involved.

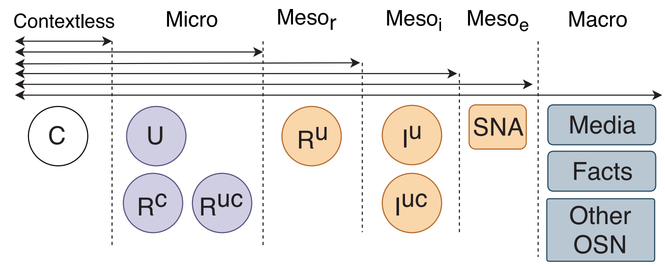

Fig. 3. List of Social Context features available at each level of analysis.

two differences are: 1) a level of analysis is added to account for analysis without social context, and 2) the meso level is further divided into three sub-levels ($meso_r$, $meso_i$, and $meso_e$), to better capture the nuances at the meso level. The result is shown in Fig. 2, and the levels are:

- Contextless: The approaches in this category do not use social context, and they rely solely on textual features.
- Micro: These approaches exploit the relation of content to its author(s), and may include other content by the same author. For instance, they may use the sentiment of previous posts [1] or other personal information such as gender and age to use a language model that better fits the user [88].
- Meso-relations ($Meso_r$): In this category, the elements from the micro category are used together with relations between users. This new information can be used to create a network of users. The slow-changing nature of relations makes the network very stable. The network can be used in two ways. First, to calculate user and content metrics, which can later be used as features in a classifier. e.g., a useful metric could be the ratio of positive neighboring users [1]. Second, the network can be actively used in the classification, with approaches such as label propagation [80].
- Meso-interactions ($Meso_i$): This category also models and utilizes interactions. Interactions can be used in conjunction with relations to create a single network or be treated individually to obtain several independent networks. The resulting network is much richer than the previous category, but also subject to change. In contrast to relations, interactions are more varied and numerous. To prevent interactions from becoming noisy, they are typically filtered. For instance, two users may only be connected only when there have been a certain number of interactions between them.
- Meso-enriched ($Meso_e$): A natural step further from $Meso_i$, this category uses additional information inferred from the social network. A common technique in this area is community detection. Community partitions may inform a classifier, influence the features used for each instance [87], or be used to process groups of users differently [22]. Other examples would be metrics such as modularity and betweenness, which can be thought of as proxies for importance or influence. Some works have successfully explored the relationship between these metrics and user behavior, in order to model users. However, these results are seldom used in classification tasks.
- Macro: At this level, information from other sources outside the social network is incorporated. For instance, Li et al. [48] use public opposition of political candidates in combination with social theories to improve sentiment classification. Another example of external information is facts such as the population of a country, or current government, which can be combined with geo-location information in social media content. A more complex example would be events in the real world or in other types of media, such as television, which can be analyzed in combination with social media activity [34].

The six levels of approaches are listed in increasing order of detail, measured as the number of elements social context may include. The specific elements that are available at each level are represented in Fig. 3. The essential elements have already been covered in the definition of social context: content (C), users (U), relations (R^c, R^u and R^{uc}), and interactions (I^u and I^{uc}). Social Context can also be enriched through SNA with techniques such as community detection (CD). Additionally, external sources of information can be used at a macro level, such as facts or hyperlinks to external media, which are not part of the definition of Social Context.

4.3. Dynamic approaches

Social context can be represented and analyzed as static or dynamic, as mentioned in the definition. Static approaches present a quasi-static view of social context and do not take its evolution into account. Note that this does not prevent context from being updated at a later point. For instance, a user label may be changed, or more content may be added. However, these changes are not integrated into the model. In most of the works analyzed, context is modeled as static. Conversely, dynamic approaches both use and need a dynamic social context, as they exploit the changing nature of social networks. These changes are an intrinsic part of the analysis and need to be part of the model.

Although none of the surveyed works use dynamic social contexts for sentiment classification, several works use dynamic social context in tasks related to sentiment analysis. Based on those and related works, we suggest dynamic approaches for sentiment analysis may adhere to the following taxonomy, depending on the parts of social context that are dynamic.

At the *Micro-dynamic* level, content is dynamic, and the changes in its activity are taken into consideration. These changes could be the increase in some metrics such as retweets and likes. For instance, the evolution in content activity (number of retweets and mentions) can be used to classify content [96].

At the *Meso-dynamic* level, inter-personal communication starts to be apparent and available. Several elements of the context can be studied in a dynamic fashion. Two types of approaches could be considered, to subdivide this level.

First, approaches that focus on virality, and are content-centric. They use the evolution of interactions, and the links between users in the network, to measure and predict future activity, or to classify content according to the activity related to it. This classification may be useful for sentiment analysis. For instance, previous works have shown different types of content are linked to different temporal patterns [96]. And by using certain features of content and its activity, it is also possible to predict further spreading in the network (i.e., a cascade) [17]. These content cascades are also linked to specific sentiments [2]. Garas et al. [26] could be relevant in this area, as it studies emotion persistence in online communications (IRC).

Second, contagion-based approaches, which are user-centric. They focus on user sentiment and emotion, instead of content. They apply social theories and experimental results regarding sentiment and emotion contagion [35]. For instance, a massive experiment on Facebook showed that emotional states can be transferred to others via emotional contagion, leading people to experience the same emotions without their awareness [45]. Hence, it may be possible to improve the prediction of a user's sentiment (and their content's) by using the sentiment of the content to which she is being exposed. On the other hand, studies of social media activity regarding grassroots movements have shown that social integration, as measured through social network metrics, increases with their level of engagement and of expression of negativity [2]. This suggests a connection between the groups to which a user belongs, and the sentiment the user expresses. The connection could be exploited for user classification and, in turn, for classification of the content created by them.

4.4. Analysis methods and social theories

Lastly, works differ in the type of classification performed. The options here range from using traditional classification algorithms (e.g., random forest, SVM) or neural networks, to network-based approaches such as label propagation. However, two types of algorithms stand out from those of contextless analysis: models that directly benefit from the networked nature of context, and deep learning approaches. Several works also use a hybrid approach, where traditional techniques are combined with network techniques, either via multiple processing steps or by combining the techniques into one.

There are several ways in which algorithms could leverage the networks in social context. Firstly, some algorithms are already network-oriented. Label propagation, in particular, has shown promising results [80], and it can be made to treat lexical resources and the subject of the analysis equally. Secondly, the structure of the network can be directly incorporated into the learning process through modified cost functions [38,92]. Thirdly, the output of a classifier can be later complemented with a network-based algorithm. For example, Li et al. [48] apply standard classification, then tweets or users are clustered, and within each cluster, every piece of content or every user are given the same label according to different criteria (i.e., most confident result, majority label, and weighted majority). Fourthly, a multi-step or ensemble classification strategy can be used, where the structure of the network and social theories are used to combine the results of different classifiers.

On the deep learning front, recent works are incorporating different types of neural networks that have been used for contextless analysis and subjectivity analysis [14], such as convolutional neural networks (CNN). At the same time, concepts such as word embeddings have inspired network embedding as an alternative way of including features from social context in the analysis [97]. The range of features that can be captured through network embeddings is vast, including several types of relations [13]. Moreover, new research is complementing and extending node embedding (i.e., nodes are represented as vectors) with other methods such as edge and community embedding [10]. In particular, community embedding has shown promising results in community prediction and node classification [12].

In general, network approaches usually follow well-known social theories. Social theories usually model how users with different views or status arrange themselves in the network. In other words, they are rules of attachment. They may also model how users behave.

Some examples of social theories or attributes include homophily, consistency, social balance, and status theory. Homophily [53] is one of the commonly used theories in the works we have examined and in the social sciences. In simple terms, homophily means a connection between two people is more likely when they are similar in some aspects (i.e., birds of a feather flock together). Under the hypothesis of homophily, when two users are connected, certain features can be propagated. Consistency [50] usually means that users tend to maintain their views over time. So, two pieces of content shared by the same user in a short period are likely to express a similar sentiment or opinion if they are about the same topic. The social status theory [47] models the balance of power in social networks. It states that, if three nodes *A, B* and *C* form a clique, and the status relation between *A* and *B* is the same as between *B* and *C*, it must also be true of *A* and *C*. In other words, the superior of your superior is your superior, and the inferior of your inferior is your inferior. Social balance models the balance of opinions in cliques. The rules in social balance translate to: a friend of a friend is a friend, and an enemy of my enemy is my friend. Tang et al. [84] presents a more detailed explanation of social theories that can be used to mine social media.

5. Review of social context and sentiment analysis works

This section is the result of reviewing the state of the art in using social context for sentiment analysis. The review is composed of five subsections. The first one presents and compares the different works that have been reviewed. The second subsection describes and compares the datasets that have been used in these works. The third subsection covers common social context features that are useful for sentiment analysis. The fourth one presents a performance comparison of the works on different datasets. The last subsection discusses ways in which sentiment analysis has been used to improve social network analysis.

5.1. Works

This section introduces recent works in the area of sentiment analysis that use social context. The aim is to compare how social context is defined and exploited in each of them. The main features of each of the works are summarized in Table 2. The table shows the gradual introduction of interactions to complement interactions, as works evolve from $meso_r$ to $meso_i$ and $meso_e$ approaches. It also highlights the most commonly used types of elements and social theories used.

To the best of our knowledge, the first work to make explicit mention of social context in the context of sentiment analysis is Lu et al. [50]. Their goal was to predict the quality of reviews, rather than their sentiment, but the work is worth mentioning for three reasons. First of all, they provide the first formal mention of social context in the sense covered in this work. Secondly, their novelty is that they merge traditional features (text) with what they call *Social Network Features*. They provide a categorization of features, including author and social network features, which are calculated with social network analysis. Lastly, the network is used to extract constraints based on several hypotheses of consistency (of authors, links, citations, and trust).

On a related note, Pennacchiotti and Popescu [67] leverage replies, retweets and friendship relations to infer user attributes, such as ethnicity and political orientation. Their definition of political orientation can be considered stance detection. Although their work is implicitly motivated by a hypothesis of homophily, they do not make any mention of specific social theories, and no constraints or rules based on them are constructed. Instead, classification is achieved via Gradient Boosted Decision Trees.

Speriosu et al. [80] introduce an alternative approach to infer polarity that exploits the networked nature of social context. They compare three different approaches: a lexicon-based classifier (baseline), a maximum entropy classifier and Label Propagation (LPROP). The best results were achieved with LPROP, which is also appealing because it yields annotations for resources (e.g., lexicon), content and users indistinctly.

Similarly, Tan et al. [83] use a network approach based on SampleRank with a Markovian model. The model assumes that the sentiment of a given user is only influenced by the sentiment label of tweets generated by that user (consistency), and the sentiment of neighboring users (homophily).

Li et al. [48] compare an approach based on linguistic features with a combination of linguistic features and social features (referred to as global social evidence). The goal is sentiment analysis about political figures (targets) on Twitter and fora. In their hybrid approach, users, targets and issues (topics targets are vocal about) form a network. Three different hypotheses are then exploited on the data: 1) global consistency on indicative target-issue pairs, 2) global consistency on indicative target-target pairs, and 3) social balance. The results are slightly better than the baseline in the case of Twitter and widely better for forum data. A similar comparison of linguistic and social features is made

Table 2
Comparison of works using sentiment analysis and social context. The number of polarity labels is shown in parentheses.

Reference	OSN	Level	l^u	l^c	i^u.	i^{uc}	r^c	$r^{u,c}$	r^u	Social Theories
Pennacchiotti and Popescu [67]	Twitter	$meso_i$	Political orientation, ethnicity	Polarity (3)		Replies, Retweets	Retweet	Authorship	Friends	
Speriosu et al. [80]	Twitter	$meso_r$	Polarity (2)	Polarity (2)				Authorship	Follower	
Tan et al. [83]	Twitter	$meso_i$	Polarity (2)	-	(mutual) mention			Authorship	Follower	consistency, homophily
Li et al. [48]	Twitter, Fora	$meso_r$, Macro	Stance (targets)	Polarity (2)	Stance (targets)					balance, consistency
Aisopos et al. [1]	Twitter	*micro*, $meso_i$		Polarity (2)	Mention			Authorship	Follower	
Hu et al. [38]	Twitter	$meso_r$	Polarity (3)	Polarity (3)				Authorship	Follower	consistency and contagion
Pozzi et al. [70]	Twitter	$meso_i$	Polarity (2)		Retweet	Retweet		Authorship	Mutual follower	
Ren and Wu [72]	Twitter	$meso_r$	Polarity (2)							homophily
Deng et al. [23]	Fora	$meso_r$		Polarity (3)			Reply		Friends, inferred friends	homophily, consistency
West et al. [92]	Wiki	$meso_i$	Polarity (3)	Polarity (3)	Votes, Mentions			Authorship		social status, social balance
Yang and Eisenstein [97]	Twitter	$meso_i$		Polarity (2)	Retweet, Mention		Retweet		Follow	language homophily
Cheng et al. [16]	Reddit	$meso_i$		Polarity (2)		Reply				
Sixto et al. [79]	Twitter	$meso_i$		Polarity (5)		Retweet		Favorite	Follow	
Xiaomei et al. [95]	Twitter	$meso_e$		Polarity (2)				Authorship	Follow	emotion contagion

by Aisopos et al. [1]. In their work, several classification algorithms are compared using different feature models, some of which include social context features.

Hu et al. [38] are the first in our review to include a classification algorithm specially tuned to incorporate social context. Their work is also interesting because they overcome the fact that most existing datasets only contain texts, which makes them unsuitable for social context analysis. They do so by combining text datasets with the friendship graph extracted from Kwak et al. [46].

Other works focus on user classification, such as Pozzi et al. [70]. They leverage connections in the network to infer user polarity, with highly positive results. User connections can also be exploited for content polarity classification. Ren and Wu [72] use both friendship and user-topic relations (calculated from user tweets) to calculate user-topic polarity. In addition to friendship, Deng et al. [23] use reply-to relations in online fora, as well as inferred friendship. West et al. [92] showed that the assumption of homophily in networks can improve polarity detection from short texts. They use social ties to infer the stance of users in Wikipedia. In particular, they exploit the social balance and social status theories. They also point out the effect that the selection strategy of training and testing nodes has on accuracy. Tang et al. [84] use similar social theories to improve sentiment analysis on Twitter.

Lately, some works have introduced novel approaches such as Convolutional Networks [97]. In doing so, they add new types of features such as network embeddings, i.e., a vector representation of the network of a user, which can be fed into a classifier. The motivation behind these embeddings is to leverage language homophily in the analysis. Cheng et al. [16] follow in these steps, with a similar premise using content from a different social network (Reddit). In this case, the analysis also exploits the fact that comments are nested at different levels.

5.2. Datasets

The usual drawback with sentiment analysis datasets is that they rarely incorporate social context. This is either because social context was not taken into consideration when the dataset was collected or because of data protection policies and terms of use of the original OSN. The latter is usually easier to circumvent, as these datasets usually have IDs or pointers to the original resources, so that the necessary data can be recovered with the appropriate credentials and access to the OSN. This process is known as hydration, and it can be used to recover more data than was initially considered. i.e., it enables the expansion of the social context. The limitation is the fact that resources can be removed or made private before hydration. Table 3 shows basic statistics of the datasets used in the works reviewed.

Table 3
Datasets used in the experiments.

	Source	Users	Entries
RT Mind [70]	Twitter	62	159
OMD [77]	Twitter	679	1261
HCR-DEV [80]	Twitter	806	1434
HCR-TEST [80]	Twitter	806	1434
STS [31]	Twitter	498	490
PF1901 [23]	Forum	412	1901
MF1560 [23]	Forum	320	1560
SemEval 2013 [56]	Twitter	3813	3813
SemEval 2014 [76]	Twitter	5749	5749
SemEval 2015 [75]	Twitter	2379	2379
Ciao [85]	Ciao	257,682	10,569
TASS [74]	Twitter	158	68,017
YANG2011 [96]	Twitter	20 M	476 M
Li-Twitter [48]	Twitter	?	4646
Li-Forum [48]	Forum	?	762
AskMen [16]	Reddit	?	1057 K
AskWomen [16]	Reddit	?	814 K
Politics [16]	Reddit	?	2180 K

RT Mind [70] contains a set of 62 users and 159 tweets, with positive or negative annotations. To collect this dataset, Pozzi et al. [70] crawled 2500 Twitter users who tweeted about Obama during two days in May 2013. For each user, their recent tweets (up to 3200, the limit of the API) were collected. At that point, only users that tweeted at least 50 times about Obama were considered. The tweets from those users that relate to Obama were kept and manually labeled by 3 annotators. The dataset contains ID of the tweet, ID of the author, text of the tweet, creation time, and sentiment (positive or negative).

The OMD dataset (Obama-McCain debate) [77] contains tweets about the televised debate between Senator John McCain, and then-Senator Barack Obama. The tweets were detected by following three hashtags: *#current,#tweetdebate*, and *#debate08*. The dataset contains tweets captured during the 97-minute debate, and 53 after it, to a total of 2.5 hours. There were 3238 tweets from 1160 people. There were 1824 tweets from 647 people during the actual debate and 1414 tweets from 738 people after it. Of those, only 1261 tweets, from 679 users, have sentiment annotations. The dataset includes tweet IDs, publication date, text, author name and nickname, and individual annotations of up to 7 annotators.

The Health Care Reform (HCR) [80] dataset contains tweets about the run-up to the signing of the health care bill in the USA on March 23, 2010. It was collected using the *#hcr* hashtag, from early 2010. A subset of the collected tweets were annotated with polarity (positive, negative, neutral and irrelevant) and polarity targets (health care reform, Obama, Democrats, Republicans, Tea Party, conservatives, liberals, and Stupak) by Speriosu et al. [80]. The tweets were separated into training, dev (HCR-DEV) and test (HCR-TEST) sets. The dataset contains tweet ID, user ID and username, text of the tweet, sentiment, target of the sentiment, annotator and annotator ID.

The Stanford Twitter Sentiment (STS) [31] contains manually annotated tweets that mention a wide range of topics such as consumer products (40d, 50d, kindle2), companies (aig, at&t), and people (Bobby Flay, Warren Buffet). The version of the dataset used by Speriosu et al. [80] contains only 216 annotated tweets, 108 of which tweets are positive, and 75 are negative. However, the original paper [31] mentions 359 tweets with positive or negative sentiment. These figures are aligned with the content of the dataset at the authors' website[1], which also includes neutral tweets, to a total of 498 tweets by 490 authors. The discrepancy should be noted, both because Speriosu et al. [80] use the reduced dataset, and because they have released a collection of three datasets together with the source code they used to process it[2]. The collection is well documented, which might make it easier for other researchers to reuse their reduced dataset.

In their work, Deng et al. [23] include two datasets. The first dataset (PF1901) is crawled from the “Election & Campaigns” board of a political forum[3], There are 1901 labeled posts in total written by 232 unique users from March 2011 to April 2012. Out of those, 419 positive and 553 negative posts are also labeled with associated candidates. The rest are considered neutral or unsure. The second dataset (MF1560) is crawled from a military forum[4], containing 43 483 threads and 1 343 427 posts. In total, there are 1560 labeled posts written by 320 unique users, out of which 437 positive and 618 negative posts also had their topic labeled. The rest are considered neutral or unsure.

The collection of SemEval datasets originate from the competition set up for the different editions of the International Workshop on Semantic Evaluation (SemEval). SemEval includes several individual tasks, which focus on different types of classification, on different types of data. For this paper, we focus on the Tweet sentiment classification tasks. There is a dataset for each edition: SemEval 2013 [56], SemEval 2014 [76], SemEval 2015 [75]. For each tweet, the dataset contains the ID of the tweet, the ID of the author, and the sentiment label (positive, negative or neutral). To use the dataset, participants are encouraged to hydrate it, using the tools provided by the organizers of the competition.

The General Corpus TASS dataset is one of the three datasets created for the *Taller de anÃ!'lisis de sentimientos* (workshop on sentiment analysis) [74]. The other two datasets are the SocialTV dataset and the STOMPOL dataset, and they are focused on aspect based analysis. The dataset contains tweets in Spanish, authored by 150 well-known personalities and celebrities of the world of politics, economy, communication, mass media and culture. The original corpus is released in XML format, and it includes date, author and ID of each tweet.

The AskMen, AskWomen and Politics datasets Cheng et al. [16][5] contain posts from popular subreddits (subcategories within the Reddit OSN[6] with different topics and styles: AskWomen (814K comments), AskMen (1057K comments), and Politics (2180K comments).

Yang and Leskovec [96] collected a dataset of nearly 476 million Twitter posts from 20 million users covering eight months, from June 2009 to February 2010. Aisopos et al. [1] filter the dataset in their work down to 6.12 million negative and 14.12 million positive tweets using emoticons. From those tweets, they finally used a sample of 1 million tweets with each polarity.

Li et al. [48] collected datasets from two OSN: an online forum and Twitter. The forum dataset was collected from the most recent posts at the “Elections & Campaigns” forum (similarly to Deng et al. [23]), from March 2011 to December 2011. 97.3% of those posts subjective, i.e., they contain positive or negative sentiments. The tweet data set was automatically collected by retrieving positive instances with *#Obama2012* or *#GOP2012* hashtags, and negative instances with *#Obamafail* or *#GOPfail* hashtags. All tweets where the hashtags of interest were not located at the very end of the message were filtered.

Lastly, the Ciao dataset [85] includes opinions on the Ciao website[7] in May 2011. The authors started the collection of the dataset with a set of most active users and then did a breadth-first search until no new users could be found. The sentiment in the dataset is expressed with a 5-star rating system.

5.3. Features

This section briefly covers some of the features that can be extracted from social context at different levels.

5.3.1. Micro features

At the micro level, features may be related to the content author, or to the content itself. From the user, the main set of features is:

- Number of followees. In OSN such as Twitter, users (followers) are exposed only to the content of their followees. This is typically an asymmetrical relation. Following another user does not require the followee to accept, except for private accounts and blocked users. For this reason, it is typical for users to follow hundreds or even thousands of users [46]. Hence, this feature is rather noisy. Some works refer to followees as friends, whereas other works reserve the term friend for mutual followers.
- Number of followers. In contrast with the previous feature, only a fraction of users tend to accumulate most of the followers [46]. As a result, the number of followers is more informative.
- Number of friends. In some instances, the number of followers that the user follows back is known. Otherwise, it has to be calculated from the meso network.
- Ratio of positive / negative / neutral content (per topic). This may indicate the typical sentiment polarity for a user. Some theories such as author coherence indicate that the sentiment we show about a topic tends to be stable over short periods. Moreover, studies show that different types of users exhibit characteristic sentiment patterns in their posts. Namely, popular users are more likely to post positive content.
- Age, gender and nationality. All these features influence the way we communicate, from the language we use to the sentiment we

[1] http://cs.stanford.edu/people/alecmgo/trainingandtestdata.zip
[2] https://bitbucket.org/speriosu/updown/
[3] http://www.politicalforum.com/elections-campaigns/
[4] http://forums.military.com/
[5] https://github.com/hao-cheng/factored_neural/
[6] https://reddit.com
[7] http://www.ciao.co.uk

are more likely to express, and they have been shown to help in sentiment analysis [88].

Content may also be linked to features such as:

- Number of favorites, retweets, and replies. These values gradually increase as more users interact with the content. For this reason, it may take some time for them to stabilize or become meaningful, and it is not available in online analysis unless some delay is added. By using specific time windows, it is also possible to snapshot the value of the metric at different times, to create derived metrics. e.g., number of replies during the first hour, and number of replies during the first day. This type of analysis also borders dynamic social context, which we have discussed earlier.
- Topic(s). The topic could either be extracted from content and metadata such as hashtags or automatically inferred with topic detection.
- Sentiment of the original message. It is only available for replies. It may be beneficial to know the original creator and the views of the creators, as that enables the use of social theories (e.g., Li et al. [48]).
- Sentiment ratio of replies. This information is not typically used because it requires a posteriori knowledge. However, for some types of offline classification, this information is known at the time of prediction.

Additionally, it is also possible to generate user and topic-specific models or to embed the context of the topical context of the content [16,23]. Network-based algorithms such as label propagation and algorithms that take arbitrary input sizes, such as recurrent neural networks, are not constrained by a fixed input space. As a result, they can incorporate features of the context without aggregation, such as averaging.

5.3.2. $Meso_r$ features

At this level, a network of users and content also starts to form. Connections in this network may be directed or undirected. Some examples of relations that can originate a network are:

- Follower relation (directed). This is the relation that, when aggregated, gives rise to the number of followees and number of followers in the previous section. It is the most common type of relation, and it typically requires further filtering, given both the tendency of users to follow hundreds of users and the lack of confirmation from the other side.
- Mutual follower relation (undirected). A simple follower relation often yields poor results. The cause could be that this type of relation is too weak [20], and is non-reciprocal. Most works use mutual relations instead, where users are only connected if they follow each other.
- Ratio of Common Followers/followees relation (undirected). This is a measure of how many followers/followees two users have in common. Under the hypothesis of homophily, it may be a proxy for user similarity. More elaborate versions may take into account the number of followees/followers of the followers/followees, via a weighted sum.
- Ratio of Common Topics/Keywords relation (undirected). Similar to the ratio of Common Followers/followees, it is related to the similarity of two users, based on the content they share.

5.3.3. $Meso_i$ features

Interactions can also be used to create a network. For instance:

- Reply interaction (directed). The act of replying forms one relation between the original content, and the content to which it replies. However, two interaction links can be formed as well: one between both users, and another one between the user and the original content. Since replies are less likely to occur than retweets, they tend to be more informative.
- Mention interaction (directed). When a user mentions another user in their content, two links are formed: a mention interaction between the two users, and a relation between the content and the user that was mentioned.
- Like/favorite interaction (directed). In most OSN, users can mark content they like. As opposed to a reply, liking is usually achieved with a single click. Hence, this is amongst the most common types of interactions.
- Retweet/reshare interaction (directed). Retweeting is the act of sharing content from a different user verbatim.
- Shared a conversation (undirected). When two users engage in a conversation (a series of replies), it can be encoded as a new interaction between the users.

The ability to relate an author to other users enables the propagation of micro features over the meso network, which yields a new set of features, such as:

- Sentiment ratio of neighbors. The ratio of positive/negative/neutral neighbors. Neighbors could be adjacent users (those sharing an edge), or users that belong to the same group (e.g., the same community). These neighbors could be filtered, e.g., to only take new neighbors into account, or neighbors that have had recent activity. The sentiment for each neighbor could also be calculated in time windows or weighted so that recent content is more important.
- Sentiment ratio of content by neighbors. Similar to the previous one, without aggregating on the user level.

Lastly, some techniques allow embedding large information networks (be it content, user or mixed networks) into low-dimensional vector spaces. These types of techniques are increasingly popular in contextless analysis due to their excellent performance [3]. The components of the embedding can then be used as features, either on their own or combined with other features. One example of network embedding is the LINE method [86], which is used in one of the works reviewed [16]. However, LINE does not take different types of nodes or relationships into account. The heterogeneous network embedding model [13] is an alternative. Although it was conceived to embed networks of text and images, it could be adapted to encode mixed networks of content and users.

5.3.4. $Meso_e$ features and enrichment through social network analysis

Social Network Analysis provides several methods to process, examine and describe a social network. These methods use the network topology and its attributes and infer information that could be useful for sentiment analysis tasks. For instance, there are several ways to measure user popularity and influence in a social network, according to different criteria. As a result, the impact of each user in the sentiment prediction can be weighted. Similarly, the importance of user connections (relations and interactions) can be measured. Thus, the granularity can be set at the connection level, where sentiment prediction is not only influenced by neighboring users, but also on the strength of the connection to those neighbors. Another example is community detection, which could help segment the user base into smaller groups that exhibit similar behavior.

5.3.5. Macro features

Macro features include any type of information that is outside of the realm of the OSN. Hence, the possibilities for features in this category are unlimited. Of all the works we have reviewed, only one [48] uses macro features. In particular, it uses known enmity or opposition between politicians, together with social theories about user and target consistency. Other possibilities include the analysis of links to external sources or attachments.

Table 4
Maximum Accuracy score reported in each work, per level of analysis and dataset.

Work	Level Dataset	Metric	Baseline	micro	$meso_r$	$meso_i$	$meso_e$	macro
[1]	YANG2011	Acc.	97.42	60.40	-	80.08	-	-
[23]	MF1560	Acc.	46.64	-	55.60	-	-	-
	PF1901	Acc.	61.24	-	72.75	-	-	-
[48]	Li-Forum	Acc.	59.61	67.24	62.89	-	-	71.97
	Li-Twitter	Acc.	83.97	-	85.35	-	-	-
[79]	TASS	Acc.	79.30	-	-	89.80	-	-
[80]	HCR-DEV	Acc.	58.60	65.70	65.20	-	-	-
	HCR-TEST	Acc.	62.90	71.20	71.00	-	-	-
	OMD	Acc.	61.30	66.70	66.50	-	-	-
	STS	Acc.	83.10	84.70	84.70	-	-	-
[95]	HCR	Acc.	69.00	-	-	-	77.5	-
	OMD	Acc.	76.00	-	-	-	76.0	-
[16]	AskMen	F1	51.70	-	-	52.70	-	-
	AskWomen	F1	55.20	-	-	56.30	-	-
	Politics	F1	53.00	-	-	54.80	-	-
[79]	TASS	F1	69.20	-	-	90.20	-	-
[97]	Ciao	F1	-	-	-	80.19	-	-
	SE 2013	F1	69.31	-	71.49	71.91	-	-
	SE 2014	F1	72.73	-	74.17	75.07	-	-
	SE 2015	F1	63.24	-	66.00	66.75	-	-

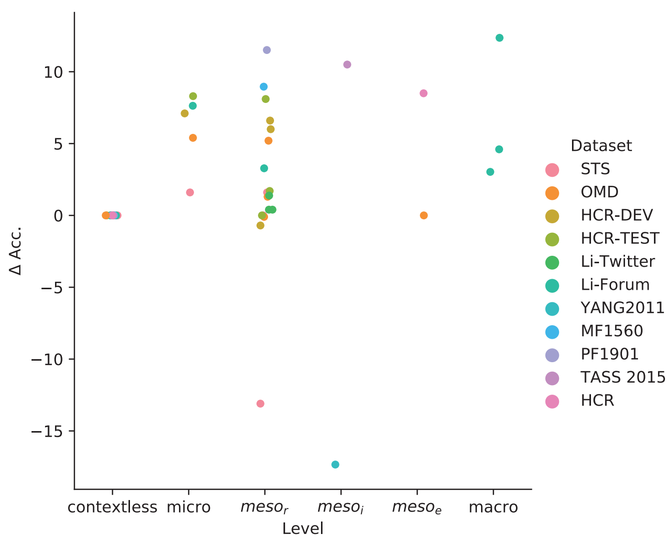

Fig. 4. Difference in accuracy with respect to a contextless approach in all works analyzed, per dataset. The results for [1] have been removed due to their unusually high accuracy (Table 4).

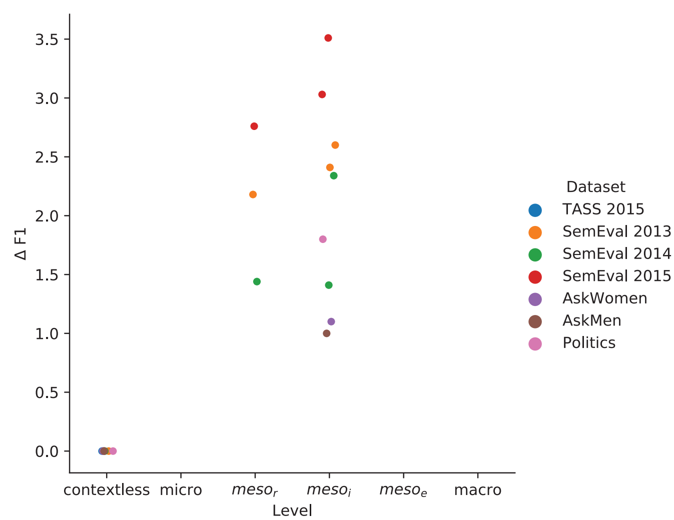

Fig. 5. Difference in F1 score with respect to a contextless approach in all works analyzed, per dataset.

5.4. Performance

Having described these works, it is also important to compare their performance. Few works use the same dataset in the same conditions. Instead of providing that comparison, Table 4 summarizes the best results for content-level classification in every work surveyed, at every level of analysis identified in the taxonomy in Section 4. The table shows both results for F1-score and accuracy, when available. As expected, the results show that social context improves the performance over the contextless baseline.

For completeness, Figs. 4 and 5 show all the results reported in these works, grouped by the level of analysis. The performance is shown relative to the contextless baseline in every dataset.

5.5. Other approaches

Although this paper focuses on using social context to improve sentiment analysis, there are other ways in which sentiment information can be fused with other sources or types of information [4]. For instance, sentiment information can be included into existing social network analysis. This can be done to characterize or explain a given phenomenon. When adding sentiment information, some patterns and trends emerge, which would otherwise be lost in the global aggregate. For instance, sentiment information can be used to analyze different Twitter communities separately instead of aggregating their results [22].

Sentiment and social network analysis can also be combined to find potentially radicalized users [6], or to highlight emotionally charged content [24]. Additionally, sentiment information alone has proved to yield very high precision and a low recall in some user classification tasks [67]. This suggests that sentiment information could be crucial in positively identifying members of specific groups.

6. Conclusions and future work

The question that motivated this work was whether there is valuable information in social networks that has the potential to improve sentiment analysis in specific scenarios. We refer to this information as social context. To answer this question, three related questions need to be answered: "what is social context?"(Q1), "can social context improve sentiment analysis?"(Q2) and "what elements of social context are more relevant for sentiment analysis?"(Q3).

To answer the first question (Q1), we analyzed the use and definitions of social context in the state of the art. Our analysis revealed that there are commonalities between these works, despite differences in notation. We formalized these commonalities in a formal definition of social context. This definition enables a richer and more precise description of social media information.

We used this definition in a new framework for comparison of approaches to sentiment analysis using social context. Part of this framework is a taxonomy of approaches, which shows the different levels of social context that are possible. Using this taxonomy, we compared works in the literature. The results of this comparison, which are included in this work, support the notion that using social context may improve performance in sentiment analysis (Q2), both in content classification and user classification tasks.

Once these levels of analysis have been identified, the natural question is what performance gains can be achieved by using more complex features. Directly comparing their results is not straightforward, but the taxonomy can be used to group approaches and to compare these groups. Higher results correspond to more detailed definitions of Social Context, as shown by $meso_i$ approaches outperforming $meso_r$ ones in most works (Q3). The trend seems to support these results, as recent works are starting to incorporate $meso_i$ approaches. Unfortunately, the number of works in the field is not enough to provide an accurate evaluation of the specific elements of content (e.g., whether retweet interactions are more informative than community detection).

On the other hand, the trend suggests that there is room for improvement in the processing of social context and its use with different classifiers. For instance, techniques such as network embeddings could be used to condense several aspects of social context.

We expect that the formal definition of context and the framework in this work foster the use of social context in sentiment analysis in two ways. Firstly, by providing a common language to express social context. Secondly, by allowing future works to perform a more systematic comparison with existing approaches. As more works start leveraging social context, the taxonomy of approaches will likely grow and add novel ideas. Similarly, more elements may need to be included in the definition of social context to account for more complex scenarios.

Acknowledgments

This work is supported by the Spanish Ministry of Economy and Competitiveness under the R&D project SEMOLA (TEC2015-68284-R) and the European Union under the project Trivalent (H2020 Action Grant No. 740934, SEC-06-FCT-2016). The authors also want to mention earlier work that contributed to the results in this paper. More specifically, the MixedEmotions (European Union's Horizon 2020 Programme research and innovation programme under grant agreements No. 644632) and SoMeDi (ITEA3 16011) projects.

References

[1] F. Aisopos, G. Papadakis, K. Tserpes, T. Varvarigou, Content vs. context for sentiment analysis: a comparative analysis over microblogs, in: Proceedings of the 23rd ACM conference on Hypertext and social media, ACM, 2012, pp. 187–196.
[2] R. Alvarez, D. Garcia, Y. Moreno, F. Schweitzer, Sentiment cascades in the 15m movement, EPJ Data Sci. 4 (1) (2015).
[3] O. Araque, I. Corcuera-Platas, J.F. Sánchez-Rada, C.A. Iglesias, Enhancing deep learning sentiment analysis with ensemble techniques in social applications, Expert Syst. Appl. (2017).
[4] J.A. Balazs, J.D. Velásquez, Opinion mining and information fusion: a survey, Inf. Fusion 27 (2016) 95–110.
[5] Y. Bengio, Learning deep architectures for AI, Found. Trends® Mach. Learn. 2 (1) (2009) 1–127.
[6] A. Bermingham, M. Conway, L. McInerney, N. O'Hare, A.F. Smeaton, Combining social network analysis and sentiment analysis to explore the potential for online radicalisation, in: Social Network Analysis and Mining, 2009. ASONAM'09. International Conference on Advances in, IEEE, 2009, pp. 231–236.
[7] M. Bolíbar, Macro, meso, micro: broadening the social of social network analysis with a mixed methods approach, Qual. Quan. 50 (5) (2016) 2217–2236.
[8] S.P. Borgatti, A. Mehra, D.J. Brass, G. Labianca, Network analysis in the social sciences, Science 323 (5916) (2009) 892–895.
[9] P. Buitelaar, M. Arcan, C. Iglesias, F. Sanchez-Rada, C. Strapparava, Linguistic linked data for sentiment analysis, in: Proceedings of the 2nd Workshop on Linked Data in Linguistics (LDL-2013): Representing and Linking Lexicons, Terminologies and Other Language Data, 2013, pp. 1–8. 00015
[10] H. Cai, V.W. Zheng, K.C. Chang, A comprehensive survey of graph embedding: problems, techniques, and applications, IEEE Trans. Know. Data Eng. 30 (9) (2018) 1616–1637. 00123
[11] E. Cambria, Affective computing and sentiment analysis, IEEE Intell. Syst. 31 (2) (2016) 102–107.
[12] S. Cavallari, V.W. Zheng, H. Cai, K.C.-C. Chang, E. Cambria, Learning community embedding with community detection and node embedding on graphs, in: Proceedings of the 2017 ACM on Conference on Information and Knowledge Management, ACM, 2017, pp. 377–386.
[13] S. Chang, W. Han, J. Tang, G.-J. Qi, C.C. Aggarwal, T.S. Huang, Heterogeneous network embedding via deep architectures, in: Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, 2015, pp. 119–128.
[14] I. Chaturvedi, E. Cambria, R.E. Welsch, F. Herrera, Distinguishing between facts and opinions for sentiment analysis: survey and challenges, Inf. Fusion 44 (2018) 65–77.
[15] H. Chen, J. Liu, Y. Lv, M.H. Li, M. Liu, Q. Zheng, Semi-supervised clue fusion for spammer detection in Sina Weibo, Inf. Fusion 44 (2018) 22–32.
[16] H. Cheng, H. Fang, M. Ostendorf, A factored neural network model for characterizing online discussions in vector space, in: Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, 2017, pp. 2296–2306.
[17] J. Cheng, L. Adamic, P.A. Dow, J.M. Kleinberg, J. Leskovec, Can cascades be predicted? in: Proceedings of the 23rd International Conference on World Wide Web, in: WWW '14, ACM, New York, NY, USA, 2014, pp. 925–936.
[18] H. Cho, J.-S. Lee, Collaborative information seeking in intercultural computer-mediated communication groups: testing the influence of social context using social network analysis, Commun. Res. 35 (4) (2008) 548–573. 00090
[19] R. Collobert, J. Weston, L. Bottou, M. Karlen, K. Kavukcuoglu, P. Kuksa, Natural language processing (almost) from scratch, J. Mach. Learn. Res. 12 (2011) 2493–2537.
[20] D. Darmon, E. Omodei, J. Garland, Followers are not enough: a multifaceted approach to community detection in online social networks, PLOS ONE 10 (8) (2015).
[21] D. Davidov, O. Tsur, A. Rappoport, Enhanced sentiment learning using twitter hashtags and smileys, in: Proceedings of the 23rd International Conference on Computational Linguistics: Posters, Association for Computational Linguistics, 2010, pp. 241–249.
[22] W. Deitrick, W. Hu, Mutually enhancing community detection and sentiment analysis on twitter networks, J. Data Anal. Inf. Process. 01 (03) (2013) 19–29.
[23] H. Deng, J. Han, H. Li, H. Ji, H. Wang, Y. Lu, Exploring and inferring user–user pseudo-friendship for sentiment analysis with heterogeneous networks, Stat. Anal. Data Min. 7 (4) (2014) 308–321.
[24] M. Gamon, S. Basu, D. Belenko, D. Fisher, M. Hurst, A.C. König, BLEWS: using blogs to provide context for news articles., in: ICWSM, 2008, pp. 60–67.
[25] B. Gao, B. Berendt, D. Clarke, R. De Wolf, T. Peetz, J. Pierson, R. Sayaf, Interactive grouping of friends in OSN: towards online context management, in: Data Mining Workshops (ICDMW), 2012 IEEE 12th International Conference on, IEEE, 2012, pp. 555–562.
[26] A. Garas, D. Garcia, M. Skowron, F. Schweitzer, Emotional persistence in online chatting communities, Sci. Rep. 2 (2012).
[27] D. Garcia, A. Abisheva, S. Schweighofer, U. Serdült, F. Schweitzer, Ideological and temporal components of network polarization in online political participatory media: ideological and temporal components of network, Policy Internet 7 (1) (2015) 46–79.
[28] A. Garcá-Pablos, M. Cuadros Oller, G. Rigau Claramunt, A comparison of domain-based word polarity estimation using different word embeddings, in: Proceedings of the Tenth International Conference on Language Resources and Evaluation, Portoroz, Slovenia, 2016.
[29] Y. Genc, Y. Sakamoto, J. Nickerson, Discovering context: classifying tweets through a semantic transform based on wikipedia, Foundations of augmented cognition. Directing the future of adaptive systems (2011) 484–492.
[30] K. Gimpel, N. Schneider, B. O'Connor, D. Das, D. Mills, J. Eisenstein, M. Heilman, D. Yogatama, J. Flanigan, N.A. Smith, Part-of-speech tagging for twitter: annotation, features, and experiments, in: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: Short Papers - Volume 2, in: HLT '11, Association for Computational Linguistics, Stroudsburg, PA, USA, 2011, pp. 42–47.
[31] A. Go, R. Bhayani, L. Huang, Twitter sentiment classification using distant supervision, CS224N Project Report, Stanford 1 (12) (2009).
[32] W. Guo, H. Li, H. Ji, M.T. Diab, Linking tweets to news: a framework to enrich short text data in social media., in: ACL (1), 2013, pp. 239–249.
[33] B. Hajian, T. White, Modelling influence in a social network: metrics and evaluation, in: 2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing, 2011, pp. 497–500.
[34] Y.-C. Heo, J.-Y. Park, J.-Y. Kim, H.-W. Park, The emerging viewertariat in South Korea: the seoul mayoral TV debate on twitter, facebook, and blogs, Telemat. Inf. 33 (2) (2016) 570–583. 00014
[35] A.L. Hill, D.G. Rand, M.A. Nowak, N.A. Christakis, Emotions as infectious diseases in a large social network: the SISa model, Proc. R. Soc. London B (2010). rspb20101217

[36] A. Hogenboom, D. Bal, F. Frasincar, M. Bal, F. De Jong, U. Kaymak, Exploiting emoticons in polarity classification of text., J. Web Eng. 14 (1&2) (2015) 22–40. 00043
[37] D. Hovy, Demographic factors improve classification performance., in: ACL (1), 2015, pp. 752–762.
[38] X. Hu, L. Tang, J. Tang, H. Liu, Exploiting social relations for sentiment analysis in microblogging, in: Proceedings of the Sixth ACM International Conference on Web Search and Data Mining, in: WSDM '13, ACM, New York, NY, USA, 2013, pp. 537–546.
[39] B.J. Jansen, M. Zhang, K. Sobel, A. Chowdury, Twitter power : tweets as electronic word of mouth, J. Am. Soc. Inf. Sci. 60 (11) (2009) 2169–2188.
[40] F. Jiang, Y.-Q. Liu, H.-B. Luan, J.-S. Sun, X. Zhu, M. Zhang, S.-P. Ma, Microblog sentiment analysis with emoticon space model, J. Comput. Sci. Technol. 30 (5) (2015) 1120–1129. 00026
[41] A.M. Kaplan, M. Haenlein, Users of the world, unite! the challenges and opportunities of social media, Bus. Horizons 53 (1) (2010) 59–68.
[42] Y. Kim, Convolutional neural networks for sentence classification, in: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), Association for Computational Linguistics, Doha, Qatar, 2014, pp. 1746–1751.
[43] Y. Kim, D. Sohn, S.M. Choi, Cultural difference in motivations for using social network sites: a comparative study of American and Korean college students, Comput. Human Behav. 27 (1) (2011) 365–372. 00708
[44] S. Kiritchenko, X. Zhu, S.M. Mohammad, Sentiment analysis of short informal texts, J. Artif. Intell. Res. 50 (2014) 723–762.
[45] A.D. Kramer, J.E. Guillory, J.T. Hancock, Experimental evidence of massive-scale emotional contagion through social networks, Proc. Natl. Acad. Sci. (2014).
[46] H. Kwak, C. Lee, H. Park, S. Moon, What is twitter, a social network or a news media? in: Proceedings of the 19th International Conference on World Wide Web, in: WWW '10, ACM, New York, NY, USA, 2010, pp. 591–600.
[47] J. Leskovec, D. Huttenlocher, J. Kleinberg, Signed networks in social media, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, 2010, pp. 1361–1370.
[48] H. Li, Y. Chen, H. Ji, S. Muresan, D. Zheng, Combining social cognitive theories with linguistic features for multi-genre sentiment analysis., in: PACLIC, 2012, pp. 127–136.
[49] Z.C. Lipton, The mythos of model interpretability., 2016 arXiv:1606.03490.
[50] Y. Lu, P. Tsaparas, A. Ntoulas, L. Polanyi, Exploiting social context for review quality prediction, in: Proceedings of the 19th International Conference on World Wide Web, ACM, 2010, pp. 691–700.
[51] G. Marcus, Deep learning: a critical appraisal., 2018 arXiv:1801.00631.
[52] J. McCrae, D. Spohr, P. Cimiano, Linking lexical resources and ontologies on the semantic web with lemon, in: Extended Semantic Web Conference, Springer, 2011, pp. 245–259. 00210
[53] M. McPherson, L. Smith-Lovin, J.M. Cook, Birds of a feather: homophily in social networks, Ann. Rev. Sociol. 27 (1) (2001) 415–444.
[54] P. Melville, W. Gryc, R.D. Lawrence, Sentiment analysis of blogs by combining lexical knowledge with text classification, in: Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, in: KDD '09, ACM, New York, NY, USA, 2009, pp. 1275–1284.
[55] T. Mikolov, K. Chen, G. Corrado, J. Dean, Efficient estimation of word representations in vector space., 2013 arXiv:1301.3781.
[56] P. Nakov, S. Rosenthal, Z. Kozareva, V. Stoyanov, A. Ritter, T. Wilson, SemEval-2013 task 2: sentiment analysis in twitter, in: Proceedings of the 7th International Workshop on Semantic Evaluation. SemEval '13, 7, 2013, pp. 312–320. Atlanta, Georgia, USA
[57] T. Nasukawa, J. Yi, Sentiment analysis: capturing favorability using natural language processing, in: Proceedings of the 2Nd International Conference on Knowledge Capture, in: K-CAP '03, ACM, New York, NY, USA, 2003, pp. 70–77.
[58] M.-T. Nguyen, D.-V. Tran, L.-M. Nguyen, Social context summarization using user–generated content and third-party sources, Know.-Based Syst. (2017).
[59] T. Noro, T. Tokuda, Searching for relevant tweets based on topic-related user activities, J. Web Eng. 15 (3–4) (2016) 249–276.
[60] P.K. Novak, J. Smailović, B. Sluban, I. Mozetič, Sentiment of emojis, PloS One 10 (12) (2015) e0144296. 00226
[61] G.K. Orman, V. Labatut, H. Cherifi, Qualitative comparison of community detection algorithms, in: International Conference on Digital Information and Communication Technology and Its Applications, Springer, 2011, pp. 265–279.
[62] E. Otte, R. Rousseau, Social network analysis: a powerful strategy, also for the information sciences, J. Inf. Sci. 28 (6) (2002) 441–453.
[63] A. Pak, P. Paroubek, Twitter as a corpus for sentiment analysis and opinion mining., in: LREc, 10, 2010, pp. 1320–1326.
[64] B. Pang, L. Lee, Opinion mining and sentiment analysis, Found. Trends®Inf. Retr. 2 (1–2) (2008) 1–135.
[65] B. Pang, L. Lee, S. Vaithyanathan, Thumbs Up?: sentiment classification using machine learning techniques, in: Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing - Volume 10, in: EMNLP '02, Association for Computational Linguistics, Stroudsburg, PA, USA, 2002, pp. 79–86.
[66] S. Papadopoulos, Y. Kompatsiaris, A. Vakali, P. Spyridonos, Community detection in social media, Data Min. Know. Discov. 24 (3) (2012) 515–554.
[67] M. Pennacchiotti, A.-M. Popescu, A machine learning approach to twitter user classification., Icwsm 11 (1) (2011) 281–288.
[68] L. Polanyi, A. Zaenen, Contextual valence shifters, in: Computing Attitude and Affect in Text: Theory and Applications, Springer, 2006, pp. 1–10.
[69] S. Poria, E. Cambria, R. Bajpai, A. Hussain, A review of affective computing: from unimodal analysis to multimodal fusion, Inf. Fusion 37 (2017) 98–125.
[70] F.A. Pozzi, D. Maccagnola, E. Fersini, E. Messina, Enhance user-level sentiment analysis on microblogs with approval relations, in: Congress of the Italian Association for Artificial Intelligence, Springer, 2013, pp. 133–144.
[71] K. Ravi, V. Ravi, A survey on opinion mining and sentiment analysis: tasks, approaches and applications, Know.-Based Syst. 89 (Supplement C) (2015) 14–46.
[72] F. Ren, Y. Wu, Predicting user-topic opinions in twitter with social and topical context, IEEE Trans. Affect. Comput. 4 (4) (2013) 412–424.
[73] L. Rokach, Ensemble-based classifiers, Artif. Intell. Rev. 33 (1–2) (2010) 1–39.
[74] J.V. Romän, E.M. Cámara, J.G. Morera, S.M.J. Zafra, TASS 2014-The challenge of aspect-based sentiment analysis, Procesamiento del Lenguaje Natural 54 (2015) 61–68.
[75] S. Rosenthal, P. Nakov, S. Kiritchenko, S. Mohammad, A. Ritter, V. Stoyanov, Semeval-2015 task 10: sentiment analysis in twitter, in: Proceedings of the 9th international workshop on semantic evaluation (SemEval 2015), 2015, pp. 451–463.
[76] S. Rosenthal, A. Ritter, P. Nakov, V. Stoyanov, SemEval-2014 Task 9: sentiment analysis in twitter, in: Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014), Dublin, Ireland, 2014, pp. 73–80.
[77] D.A. Shamma, L. Kennedy, E.F. Churchill, Tweet the debates: understanding community annotation of uncollected sources, in: Proceedings of the First SIGMM Workshop on Social Media, in: WSM '09, ACM, New York, NY, USA, 2009, pp. 3–10.
[78] A. Sharma, S. Dey, A comparative study of feature selection and machine learning techniques for sentiment analysis, in: Proceedings of the 2012 ACM Research in Applied Computation Symposium, ACM, 2012, pp. 1–7.
[79] J. Sixto, A. Almeida, D. López-de Ipiña, Analysis of the structured information for subjectivity detection in twitter, Trans. Comput. Collect. Intell. XXIX (2018) 163–181.
[80] M. Speriosu, N. Sudan, S. Upadhyay, J. Baldridge, Twitter polarity classification with label propagation over lexical links and the follower graph, in: Proceedings of the Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, 2011, pp. 53–56.
[81] J.F. Sánchez-Rada, C.A. Iglesias, Onyx: a linked data approach to emotion representation, Inf. Process. Manag. 52 (1) (2016) 99–114. 00026
[82] M. Taboada, J. Brooke, M. Tofiloski, K. Voll, M. Stede, Lexicon-based methods for sentiment analysis, Comput. Linguist. 37 (2) (2011) 267–307.
[83] C. Tan, L. Lee, J. Tang, L. Jiang, M. Zhou, P. Li, User-level sentiment analysis incorporating social networks, in: Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, in: KDD '11, ACM, New York, NY, USA, 2011, pp. 1397–1405.
[84] J. Tang, Y. Chang, H. Liu, Mining social media with social theories: a survey, SIGKDD Explor. Newsl 15 (Iid) (2014) 20–29.
[85] J. Tang, H. Gao, H. Liu, mTrust: discerning multi-faceted trust in a connected world, in: Proceedings of the Fifth ACM International Conference on Web Search and Data Mining - WSDM '12, ACM Press, Seattle, Washington, USA, 2012, p. 93.
[86] J. Tang, M. Qu, M. Wang, M. Zhang, J. Yan, Q. Mei, Line: Large-scale information network embedding, in: Proceedings of the 24th International Conference on World Wide Web, International World Wide Web Conferences Steering Committee, 2015, pp. 1067–1077.
[87] A. Tommasel, D. Godoy, A social-aware online short-text feature selection technique for social media, Inf. Fusion 40 (2018) 1–17.
[88] S. Volkova, T. Wilson, D. Yarowsky, Exploring demographic language variations to improve multilingual sentiment analysis in social media, in: Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing, 2013, pp. 1815–1827.
[89] S. Volkova, Predicting Demographics and Affect in Social Networks, Johns Hopkins University, Baltimore, Maryland, 2015 Ph.D. thesis.
[90] S. Wang, C.D. Manning, Baselines and bigrams: simple, good sentiment and topic classification, in: Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Short Papers - Volume 2, in: ACL '12, Association for Computational Linguistics, Stroudsburg, PA, USA, 2012, pp. 90–94.
[91] W. Wei, J.A. Gulla, Sentiment learning on product reviews via sentiment ontology tree, in: Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics, Association for Computational Linguistics, 2010, pp. 404–413. 00134
[92] R. West, H.S. Paskov, J. Leskovec, C. Potts, Exploiting social network structure for person-to-person sentiment analysis, CoRR (2014) abs/1409.2450.
[93] F. Wu, J. Shu, Y. Huang, Z. Yuan, Co-detecting social spammers and spam messages in microblogging via exploiting social contexts, Neurocomputing 201 (2016) 51–65.
[94] R. Xia, C. Zong, Exploring the use of word relation features for sentiment classification, in: Proceedings of the 23rd International Conference on Computational Linguistics: Posters, in: COLING '10, Association for Computational Linguistics, Stroudsburg, PA, USA, 2010, pp. 1336–1344.
[95] Z. Xiaomei, Y. Jing, Z. Jianpei, H. Hongyu, Microblog sentiment analysis with weak dependency connections, Know.-Based Syst. 142 (2018) 170–180.
[96] J. Yang, J. Leskovec, Patterns of temporal variation in online media, in: Proceedings of the Fourth ACM International Conference on Web Search and Data Mining, ACM, 2011, pp. 177–186.
[97] Y. Yang, J. Eisenstein, Overcoming language variation in sentiment analysis with social attention, Trans. Assoc. Comput. Linguist. 5 (2017) 295–307.