Sentiment Analysis in Social Media

Georgios Paltoglou

Abstract Sentiment Analysis deals with the detection and analysis of affective content in written text. It utilizes methodologies, theories, and techniques from a diverse set of scientific domains, ranging from psychology and sociology to natural language processing and machine learning. In this chapter, we discuss the contributions of the field in social media analysis with a particular focus in online collective actions; as these actions are typically motivated and driven by intense emotional states (e.g., anger), sentiment analysis can provide unique insights into the inner workings of such phenomena throughout their life cycle. We also present the state of the art in the field and describe some of its contributions into understanding online collective behavior. Lastly, we discuss significant real-world datasets that have been successfully utilized in research and are available for scientific purposes and also present a diverse set of available tools for conducting sentiment analysis.

1 Introduction

The unprecedented rise in the popularity of social media platforms (Harvey 2010), exemplified by blogs, forums, and services such as Facebook, Twitter, Google Plus, etc. and their infiltration in everyday life (e.g., Fouche 2011; Barnett 2009) has resulted in an important paradigm shift in the way that people communicate with each other online and, more generally, interact with the Web. Previously, users were typically limited to consuming content authored by professionals, such as news agencies or corporations. In contrast, nowadays they can effortlessly create

G. Paltoglou (✉)
School of Technology, University of Wolverhampton, Wulfruna Street, Wolverhampton WV1 1LY, UK
e-mail: g.paltoglou@wlv.ac.uk

and share their own content and seamlessly interact with other users within a network of peers, often in a real-time, synchronous manner.

The importance of this phenomenon and its repercussions to society were vividly demonstrated in 2011, when a series of sociopolitical events, such as the London riots[1] and the Arab Spring,[2] took place. In both cases, online social media were regarded to significantly contribute to the emergence and proliferation of the events, with one participant of the latter event claiming that "We use Facebook to schedule the protests, Twitter to coordinate and YouTube to tell the world" (Howard 2011). Those effects were more pronounced by the fact that even official authorities considered (Halliday and Garside 2011) or took direct action (Gazzar et al. 2011) in shutting down Internet communications in order to prevent people from having access to such services.

In this chapter, we analyze the effects and implications that the new field of sentiment analysis can have in this novel environment. The next section provides a concise but thorough introduction to the field, and in Sect. 3 we discuss the applications of the field in social media in general and in the context of online collective behavior, more specifically. Sections 4 and 5 present some important, real-world datasets and tools that have been successfully used by researchers and are freely available, and in Sect. 6 we conclude and summarize.

2 Sentiment Analysis

Sentiment analysis is a subdiscipline within data mining, machine learning, and computational linguistics and also borrows elements from psychology and sociology. Generally, it deals with the computational treatment of expressions of opinion, sentiment, emotion, beliefs, and speculation in written text (Wiebe et al. 1999). Those are concisely defined as *private states*, i.e., states that are not open to objective observation or verification (Quirk et al. 1985). The field has also been referenced in research as *opinion mining* or *subjectivity analysis*,[3] and we will use those terms interchangeably in this chapter.

More specifically, opinion mining addresses the problem of detecting, extracting, analyzing, and quantifying expressions of private states in written text in an automatic, computer-mediated fashion. Particular emphasis should be placed on the term "computer-mediated," as the field has a particular focus on designing, analyzing, and implementing software that performs the aforementioned analysis in an automatic manner.

---

[1] http://en.wikipedia.org/wiki/2011_England_riots

[2] http://en.wikipedia.org/wiki/Arab_Spring

[3] For more information about the terminology in the field, we refer the interested reader to chapter 1.5 of Pang and Lee (2008).

As a result, sentiment analysis software receives as input unstructured data, such as the textual exchanges between users (e.g., tweets, Facebook updates, blog/forum posts, etc.) which by themselves are of use only to language specialists, and provides as output an informed estimate of the sentiment contained within the exchange. That estimate can take a diverse set of forms depending on a number of factors, such as the specific prerequisites of the application, the domain of interest, the psychological paradigm adopted, etc. Typical examples of valued output can include but are not limited to:

- A *binary* decision indicating whether the affective content of the communication belongs to one of two predefined categories. For example, an opinion about a new political legislation can have a positive or negative position about it, supporting or rejecting it, respectively (Whitelaw et al. 2005). In some applications, such as online discussions, where not all exchanges are necessarily affective (e.g., Chmiel et al. 2011a; Mitrović et al. 2011), a *ternary* scheme is more appropriate and adopted: {objective, positive, negative}, where the *objective* category typically signifies the absence of opinionated or affective content, such as encyclopedic-type, mainly informative content.
- A *real* value providing more fine-grained and detailed information about affective content. Typical applications can include studies on the level of *valence* or *arousal* at a specific scale (e.g., [1,9]) expressed in a forum post (Gonzalez-Bailon et al. 2010; Dodds and Danforth 2009; Paltoglou et al. 2013). Valence is defined as the dimension of experience that refers to hedonism (i.e., pleasure and displeasure) and arousal refers to the level of excitement or energy of the individual (Barrett and Russell 1999).
- A *categorical* classification where the analysis aims to determine the general psychological state of the author of a message. Typically, the analysis will involve several potential states such as nervousness, anxiety, fear, fatigue, and tension (Mishne 2005; Bollen et al. 2011). In the same manner, *basic emotions*, such as love, hate, etc. (Dalgleish and Power 1999) can be detected in written text (Strapparava and Mihalcea 2008), although there is significant debate within the field of psychology on the human agreement (Strapparava and Mihalcea 2007) and universality (Mauss and Robinson 2009) of such states.

Sentiment analysis is a nontrivial task, as even people often disagree on the affective content of written text (Paltoglou et al. 2010; Strapparava and Mihalcea 2007). Prosaic elements, such as irony and thwarted expectations (occurring when a change of opinion takes place in the end) pose particular challenges. Contextuality is also often vital; a review comprising only of the sentence "go read the book!" would be positive in a book review, but negative if referring to a movie. People also often find unique ways of expressing affect without necessarily using affective words and occasionally communicate ambiguous messages.

There are also additional issues pertaining to social media, because their typical content does not necessarily conform to the standard syntactic and grammar rules. In contrast, it contains idiomatic expressions which varies significantly based on the users' social background (Thelwall 2009), heavily utilizes acronyms and

Table 1 Examples of textual communication with affective content

Source	Text
MySpace	hey witch wat cha been up too
Twitter	It saddens my heart 2 C women with black eyes! :(
YouTube	were the f. . . are the lyrics u suck man. . .but the song is great..SOAD forever
Forums (BBC)	And I think you are a slanderous and ignorant fool. . . Great how free speech works, isn't it? . . . I love it!
Digg	This story makes me happy and sad at the same time. . .

emoticons, and is overall highly heterogeneous and often targeted to specific social groups. Table 1 presents some examples of the aforementioned challenges from a variety of social media.

2.1 Behind the Scenes

Machine Learning techniques (Chen and Zimbra 2010; Sebastiani 2002) have been an integral part of opinion mining, as a significant number of sentiment analysis solutions are based on them. According to this approach, a general inductive process initially *learns* and stores the characteristics of a category (e.g., opinions in favor of some legislation) during a training phase. This is achieved by observing the properties of a set of humanly annotated, preclassified text segments. Those preclassified text segments, which can be forum posts, political speeches, etc., comprise the *training dataset*.

Creating such a dataset is generally a time-consuming task, as it typically requires manual, human effort in order for the text segments to be read, understood, and assigned to a category. Nonetheless, there are ways in which the process can be done in an automatic or semiautomatic way, for example, by examining the metadata that accompany the textual message, such as the "number of stars" in product reviews (Pang et al. 2002) or the ideological stand or final vote in political issues (Thomas et al. 2006). Alternatively, implicit signals within the message itself, e.g., the type of emoticons used (Pak and Paroubek 2010) can be used to infer an overall affective state. Lastly, crowd sourcing techniques can provide an alternative solution to producing such annotations (Brew et al. 2010).

The knowledge that is acquired through the training phase is later *applied* to determine the best category for new, unseen text segments (Sebastiani 2002). Based on this general theoretical foundation, a number of sentiment analysis techniques have been presented that utilize specific machine learning algorithms, such as Naive Bayes (John and Langley 1995), Logistic Regression (Le Cessie and Van Houwelingen 1992), Support Vector Machines (Platt 1999; Joachims 1999), and others. A detailed discussion on machine learning is beyond the scope of this book, but we refer the interested reader to the books of Mitchell (1997) and Bishop (2006) for a thorough introduction to the topic.

Often, sentiment analysis approaches that utilize machine learning take advantage of the idiosyncrasies of affective communication and extend the standard features/properties of analyzed documents with additional, sentiment-based elements (Velikovich et al. 2010; Whitelaw et al. 2005; Agarwal et al. 2011). For example, limited human assistance can be employed in annotating specific emotionally definitive phrases (Zaidan et al. 2007) or analyzing the syntax of the text in order to extract useful patterns (Wilson et al. 2005b). More often than not, such properties can be extracted from *emotional* dictionaries; lexicons in which lemmas are annotated with affective semantics, for example, the level of positivity or negativity they typically convey.

There is a significant number of such lexicons in research, which have been produced either automatically or semiautomatically (Jijkoun et al. 2010), that usually extend the WordNet (Miller 1995)[4] lexical database with additional annotations. Examples include WordNet-Affect (Strapparava and Valitutti 2004) and SentiWordNet (Baccianella et al. 2010) which adopt a different annotation scheme. The former contains 4,787 words, mainly nouns and verbs, that directly or indirectly refer to mental states. For example, the term "anger" is annotated as referring to "emotion," while "cry" belongs to the "behavior" category. SentiWordNet, on the other hand, focuses on a simpler ternary scheme and gives each lemma three scores based on how positive, negative, or objective it is. The three scores sum up to 1, giving the annotations an interesting probabilistic interpretation. For example, the noun "love" has a positive value of 0.625 and negative value of 0.0, while "hate" has a negative value of 0.75 and a positive value of 0.0.

In addition, there are also affective lexicons that were produced by psychological studies and are manually annotated by human coders. Those include the "Linguistic Inquiry and Word Count" or LIWC (Pennebaker and Francis 1999) and the "Affective Norms for English words" or ANEW (Bradley and Lang 1999) lexicons which also offer different types of annotations. LIWC classifies words in one or several, not necessarily affective, categories, such as social, family, time, positive, anger, etc. ANEW provides for each word three values of valence, arousal, and dominance on a [1–9] scale. Both have been used in a number of large-scale studies (e.g., Owsley et al. 2006; Bollen et al. 2011; Gonzalez-Bailon et al. 2010).

Affective lexicons have also been utilized in nonmachine learning solutions to opinion mining. Typically, such approaches can be very effective in cases where training data is particularly scarce. In addition, the fact that they do not need training and can often be applied *off the shelf* is often seen as a significant advantage. They have been known to perform adequately effectively in a number of diverse environments (Paltoglou and Thelwall 2012), often reaching human-level accuracy in certain cases (Thelwall et al. 2010). Examples include the Opinion Observer (Ding et al. 2008), which is mostly aimed at estimating the polarity of product reviews, the Semantic Orientation CALculator or SO-CAL software

---

[4] WordNet is a lexical database of English words which in addition to standard definitions also provides semantic relations between words.

(Taboada et al. 2001), that provides a ternary scheme, and SentiStrength (Thelwall et al. 2010), which estimates the strength of positive and negative sentiment by producing two separate values: -1 (not negative) to -5 (extremely negative) and 1 (not positive) to 5 (extremely positive). SentiStrength was designed for application in social media exchanges, where as noted before, communication is typically short and informal.

Most lexicon-based approaches, as those discussed above, are usually based on estimating the sentiment content of textual communication by utilizing one or more affective lexicons. In addition, in order to provide increased effectiveness, they also incorporate syntactical-based capabilities, such as detecting negation and emoticons and/or typically incorporate lists of intensifier/diminisher words that increase or decrease respectively the affective strength of sentiment words. Such additional capabilities would provide them the ability to distinguish the difference of valence between phrases such as "I don't love you!" and "I love you very much!", even though both contain the same affective word "love."

3   Sentiment Analysis in Social Media

The aforementioned increase of user-generated content has resulted in creating a digital landscape where the application of opinion mining can provide invaluable information and insight about the affective state of individuals. As online participation nowadays very often accompanies offline activity, the results can provide useful insights into their general well-being and behavior (Kramer 2010). Importantly, applied in massive scale (e.g., Godbole et al. 2007; Kramer 2010), sentiment analysis can provide useful insights about groups or collectives of people.

The means through which it can produce such analyses is by processing the textual content of online social media communication and providing concrete and quantifiable information about its affective properties. Group-level analysis can be obtained by automatically analyzing the *public* communication between individual group members or their messages to the outside world and aggregating the results in order to formulate an overview of the collective affective properties of the group's communication.

The produced analysis can be of different granularity levels, depending on the specific requirements of the application and the interests of the analyzing party. For example, individuals may be grouped together by the emotional properties of their communication (Chmiel et al. 2011a), their political stand (Thomas et al. 2006), or by the discussion threads they participate in (Gonzalez-Bailon et al. 2010). Typically, it is useful to trace the textual exchanges of group members through time, effectively creating a temporal-based analysis in regard to affective communication in order to observe the progression of emotion throughout the life cycle of important events (Diakopoulos and Shamma 2010; Bautin et al. 2008; Chang et al. 2011).

As most collective actions are strongly linked to, typically negative, affective states, such as anger, indignation, or high levels of arousal (Russell 1980), such an

analysis can provide unique insights throughout the formulation, development, and death of collective phenomena. This information can be of paramount importance in understanding the intricate workings of collective action by providing evidence of specific affective states (e.g., negativity) during the life span of such actions. For example, an analysis of tweets relating to the 2011 London riots showed that Twitter was mostly used to notify users about subsequent events, rather than promote illegal activity (Tonkin et al. 2012).

Chmiel et al. (2011a) study the role of emotions in the life span of online communities. They investigate whether the affective properties of the communication amongst members can quantitatively and qualitatively influence the community's trajectory in time (i.e., whether they will dissolve quickly or persist). They cluster posts from blogs, forums, and other social media based on the similarity of their emotional valence, and one of the findings they present is that the length of such clusters can be significantly affected by their emotional properties, compared to a random clustering. Based on those results, they conclude that collective emotional states created and propagated in online communities can be of vital importance for the survival of those communities. In a similar fashion, Mitrović et al. (2011) show that strong, negative emotions play a critical part in the formulation, survival, and growth of online communities. Specifically, they demonstrate that one of the driving mechanisms of thriving communities is the, mostly negative, emotional state and communication of their users, a finding that is also confirmed by Gonzalez-Bailon and Paltoglou (2012). They, in turn, study the whole life cycle of an online community from birth to dissipation and show that negative posts and their authors tend to be more popular than positive ones. They also conduct a longitudinal analysis to show that increases of positive comments have a negative impact on the activity of the community, concluding that, potentially in contrast to popular belief, disagreement and discontent are crucial to keep communities together and alive.

Twitter, as expected, has attracted significant attention in regard to the affective properties of the exchanges of its users (Agarwal et al. 2011). Bollen et al. (2011) view tweets as "temporally-authentic microscopic instantiations of public mood state" and attempt to extract six dimensions of mood from them: tension, depression, anger, vigor, fatigue, and confusion. They discover that public mood is indeed closely correlated with wider social and economical phenomena, such as stock market and crude oil prices. Importantly, they also report that significant events in the social or political arena have direct and measurable effects in the public mood as expressed in tweets. Thelwall et al. (2011) apply SentiStrength (Thelwall et al. 2010) in Twitter and discover that popular events, regardless of their actual nature, are typically associated with increases in negative sentiment strength, confirming the aforementioned findings on the importance of negativity in online communities. Diakopoulos and Shamma (2010) develop an analytical methodology for understanding the temporal dynamics of sentiment in relation to televised events (in the particular case, presidential debates) and demonstrate visuals and metrics that can be utilized to measure aggregated group emotions, anomalies, and indications of controversial topics. On the same topic, Pennebaker (2008) and Pennebaker

and Persaud (2010) are able to provide useful insights into the psychology of the candidates by analyzing their words and the frequency with which they are used during the discussions.

Kramer (2010) examines the use of emotion-barring words for 100 million Facebook users and reports that their usage closely follows self-reported satisfaction with life and that expected sentiment peaks occur at emotionally significant celebrations (e.g., Thanksgiving for U.S. users). Gonzalez-Bailon et al. (2010) show that automatic estimations of sentiment of public opinion can predict presidential approval rates, indicating a strong connection between online sentiment and official polls. More generally, they show that political events can incite changes in emotional perceptions and that these perceptions can shape political attitudes. Similarly, O'Connor et al. (2010) show that sentiment analysis techniques can predict political opinion and consumer confidence with high correlation, showing that they provide a good estimator of public feeling.

Mishne and de Rijke (2006) capture the "blogosphere state-of-mind" by analyzing and aggregating the affective content of 3.5 million blog posts over 39 days from data extracted from LiveJournal,[5] one of the largest blogging communities. They provide two case studies, one from the 7 July 2005 London bombing and a second from a recurring event, in the particular case examining the drinking habits of bloggers during the weekend. Their analysis shows that in the former case there is a significant increase in the reporting of negative emotions, such as "sadness" and "shock" while at the same time there was a significant decrease of positive emotions. They were also able to detect the *irregular* mood behavior that resulted from the event. In the latter case, of a recurring phenomenon, they were able to detect spikes of frequency of words relating to increased alcohol consumption during the weekend periods. Overall, their results strongly indicated that changes in public mood in relation to either irregular events or recurring events can be detected in the online communication of social media users.

In conclusion, it can be seen that the application of sentiment analysis in social media communication can provide invaluable information concerning the emotional state of individuals and collectives. The discussed research has shown that emotions play a vital role in the survival of online communities and often accompany significant worldwide sociopolitical events. Through the application of opinion mining, such events can be effectively detected and their perception by the general public extracted and quantified.

---

[5] http://www.livejournal.com

4 Datasets

In this section, we introduce and briefly discuss some of the datasets that have been used in the past in research. When the discussed dataset is publicly available, we'll make a note of it.

Ideally, researchers would like to have direct access to social network sites in order to be able to conduct research with all available data. For various reasons, such as user privacy or corporate policies, that is rarely possible, especially when the site is closed, i.e., users can only access the data of limited other users that have explicitly allowed them to do so, e.g., Facebook. There are studies nonetheless (Kramer 2010; Velikovich et al. 2010) that are typically conducted by employees of such services that often provide significant insights because of the massive scale of the analysis that they are based on.

Twitter on the other hand provides significant advantages over Facebook, because everyone can access the majority of its contents, since most of it is public. In fact, as we've seen, content from the site has been used numerous times by researchers (Bollen et al. 2010, 2011; Pak and Paroubek 2010; Paltoglou and Thelwall 2012). Unfortunately, its Terms of Service[6] prohibit the distribution of any content to third parties; therefore, any collected data is usually prevented from being given to other researchers. Thankfully, the Twitter API[7] provides easy access to the site's content through programming. In a similar fashion, the YouTube API[8] also allows programs to perform most of the operations available on the actual site, such as searching for videos, retrieving feeds, seeing relevant content, etc. Therefore, even though it is not always easy to distribute data from social media, for a significant number of sites there are ways in which such data can be collected and analyzed by individual researchers.

Despite the aforementioned limitations, there is a significant number of publicly available datasets for research purposes. Paltoglou et al. (2010) present two datasets extracted from BBC discussion forums and the Digg Web site. The former includes about 2.5 million textual exchanges on 100 K discussion threads, such as politics, religion, and UK and World news, spanning four years, from June 2005, when the forum went online, until June 2009, when part of it was shut down. Digg is a social news Web site, where people link to news articles they find interesting, and other users discuss them on the Web site or vote them up or down. The specific dataset comprises a full three month crawl of all activity on the Web site, from February to April 2009 and includes about 1.2 million stories and 1.6 million individual comments. Both datasets have been extensively used in research in studies of online collective phenomena (Chmiel et al. 2011a, b; Mitrović et al. 2011). More details on

---

[6] https://twitter.com/tos

[7] API stands for "Application Protocol Interface" and usually provides methods for accessing the content of services or software through programming techniques. A guide to the Twitter API can be found here: https://dev.twitter.com/

[8] http://code.google.com/apis/youtube/overview.html

the datasets as well as information on how they can be obtained can be found in Paltoglou et al. (2010).

Another set of important datasets are the ICWSM Spinn3r Datasets (Burton et al. 2009). There are two versions of the datasets, one from 2009[9] and a more recent one from 2011.[10] Both datasets are provided by Spinn3r.com and include several million blog posts crawled by Spinn3r. The former dataset comprises of 44 million blog posts, published between 1 August and 1 October 2008, to an uncompressed size of 142 GB. It is offered in a preprocessed XML format, in order to facilitate researchers analyze its content. Its time span includes a number of significant worldwide news events, such as the Olympic Games, the U.S. presidential nominating conventions, the beginning of the credit crisis, etc. The 2011 dataset (Burton and Soboroff 2011) comprises of 386 million blog posts, new articles, forum posts, and social media communication published between 13 January and 14 February 2011, to an uncompressed size of approx. 3 TB. Its time span includes a significant number of events relating to the Arab Spring, including the Egyptians protests, the Tunisian revolution, and others. Both datasets have already been used in sentiment analysis studies and other types of research (e.g., Cha et al. 2009) and are available for research purposes.

5   Sentiment Analysis Tools

There are a significant number of machine learning tools that are freely available for research, and sometimes industrial, purposes. Most of those either provide optimized versions of specific classification algorithms or, alternatively, general frameworks incorporating several classifiers.

Typical examples of optimized classifiers include SVM^{light} (Vapnik 1999; Joachims 1999), LIBSVM (CC. Chang and Lin 2011), and LIBLINEAR (Fan et al. 2008) all of which provide different implementations of Support Vector Machine (SVM) classifiers, which are considered state of the art in terms of classification accuracy. LIBLINEAR in particular is explicitly designed for application in massive datasets with millions of documents and features, making it an ideal candidate for such environments.

General frameworks of classifiers include Weka (Hall et al. 2009) which provides an extended set of machine learning algorithms (SVM, Rule-based algorithms, etc.) for data mining tasks. In addition to classification, Weka provides tools for clustering, regression, and visualization, making it a very capable, all-around tool for analyzing data. Other tools that have been used in research and are freely available include LingPipe (Alias-i 2008), Mallet (McCallum 2002), and Apache

---

[9] Available at: http://www.icwsm.org/2009/data/index.shtml

[10] Available at: http://icwsm.org/data/index.php

OpenNLP[11] all of which are fully fledged toolkits for processing text using natural language processing and machine learning techniques. They both provide a wealth of tools for processing text, such as named-entity recognition, part-of-speech tagging, sentence segmentation, etc. (Manning and Schütze 1999).

As most of these tools implement machine learning techniques, they will typically require some sort of training in order to be applied in realistic scenarios, therefore, a *training dataset* is necessary (see Sect. 2.1). In rare cases, already trained models for certain tasks are provided with some of the tools. For example, LingPipe provides trained models for part-of-speech tagging and Chinese word segmentation,[12] but to our knowledge no trained models for sentiment analysis are available.

Lexicon-based approaches provide a potential solution to the issue, since in most cases they can by applied *off the shelf* without any training or need for human labor. Typical examples include the lexicon-based classifier by Paltoglou and Thelwall (2012), SentiStrength (Thelwall et al. 2010), and OpinionFinder (Wilson et al. 2005a). The former two[13] have been explicitly designed and tested for the type of informal textual communication that is typical in online discussions, tweets, and social network communication. The former provides a ternary classification scheme where a textual segment is classified as *{objective, positive, negative}* and the latter is mainly focused on measuring the level of positivity and negativity in text. The OpinionFinder[14] system performs various levels of affective analysis, by identifying subjective sentences and extracting various aspects in relation to them, such as the opinion holder (i.e., the entity expressing an opinion), the opinion target, etc.

6 Summary and Conclusions

In this chapter, we focused on the field of sentiment analysis and demonstrated how it can contribute to deepening our understanding of online communities and collective actions. We discussed how the research can be utilized to analyze such phenomena throughout their life cycle and provided insights into the internal mechanisms that drive them. We also presented some of the results that have already been produced, demonstrating the importance of emotions in the creation, dissemination, survival, and dissipation of online communities and collective actions. It is interesting to note that those results have come from a variety of scientific domains, e.g., sociology, complex systems and physics, indicating the

---

[11] http://incubator.apache.org/opennlp/

[12] http://alias-i.com/lingpipe/web/models.html

[13] Available at: http://www.cyberemotions.eu/data.html

[14] Available at: http://code.google.com/p/opinionfinder/

diversity of the analyses that is possible on the output of opinion mining algorithms applied to social media communication.

Additionally, we presented relevant, real-world, large-scale datasets that are publicly available for interested researchers. Some of them are focused, containing information and content from specific social media Web sites for varied periods of time, which extend from some months to several years, while others are more general and are comprised of data extracted from a wealth of sources, such as blogs, forums, and discussion boards, which extend a limited amount of time. We also discussed how despite the fact that some previously used datasets are not available for distribution due to the Terms of Services of individual social media Web sites, programming interfaces are typically provided that can give access to their contents.

Lastly, we presented the state-of-the-art in efficient and effective ways of conducting sentiment analysis in social media, using either machine-learning techniques or unsupervised, lexicon-based approaches. Importantly, we presented freely available tools that can be used for sentiment analysis tasks, either after some training, in the case of the former approaches, or *off the shelf* for the latter solutions.

References

Agarwal A, Xie B, Vovsha I, Rambow O, Passonneau R (2011) Sentiment analysis of Twitter data. In: Proceedings of LSM'11. Association for Computational Linguistics, Stroudsburg, PA, pp 30–38

Alias-i (2008) Lingpipe 4.1.0. http://alias-i.com/lingpipe

Baccianella S, Esuli A, Sebastiani F (2010) Sentiwordnet 3.0. In: Proceedings of LREC'10, Valletta, Malta

Barnett E (2009) Facebook fuelling divorce, research claims (21 Dec, 2009, The Telegraph). http://www.telegraph.co.uk/technology/facebook/6857918/Facebook-fuelling-divorce-research-claims. html. Accessed 08 Sept 2011

Barrett LF, Russell JA (1999) The structure of current affect: controversies and emerging consensus. Curr Dir Psychol Sci 8(1)

Bautin M, Vijayarenu L, Skiena S (2008) International sentiment analysis for news and blogs. In: Proceedings of ICWSM'08, Seattle, Washington, DC

Bishop CM (2006) Pattern recognition and machine learning (information science and statistics). Springer, New York

Bollen J, Mao H, Zeng XJ (2010) Twitter mood predicts the stock market. CoRR. http://arxiv.org/abs/1010.3003

Bollen J, Mao H, Pepe A (2011) Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena. In ICWSM, Barcelona, Spain

Bradley MM, Lang PJ (1999) Affective norms for English words (ANEW): instruction manual and affective ratings. The Center for Research in Psychophysiology, Gainesville, FL

Brew A, Greene D, Cunningham P (2010) Using crowdsourcing and active learning to track sentiment in online media. In: Proceedings of ECAI'10, Lisbon, Portugal, pp 145–150

Burton KNK, Soboroff I (2011) The ICWSM 2011 spinn3r dataset. ICWSM, Barcelona

Burton K, Java A, Soboroff I (2009) The ICWSM 2009 spinn3r dataset. ICWSM, San Jose, CA

Cha M, Pérez JAN, Haddadi H (2009) Flash floods and ripples: the spread of media content through the blogosphere. In: Proceedings of ICWSM'11, Barcelona, Spain

Chang CC, Lin CJ (2011) Libsvm: a library for support vector machines. ACM Trans Intell Syst Technol 2:27.1–27.27

Chang R, Pimentel S, Svistunov A (2011) Sentiment analysis of occupy wall street tweets. http://cs229.stanford.edu/proj2011/ChangPimentelSvistunov-SentimentAnalysisOfOccupyWallStreetTweets.pdf

Chen H, Zimbra D (2010) Ai and opinion mining. IEEE Intell Syst 25:74–80

Chmiel A, Sienkiewicz J, Thelwall M, Paltoglou G, Buckley K, Kappas A (2011a) Collective emotions online and their influence on community life. PLoS ONE 6(7):e22207

Chmiel A, Sobkowicz P, Sienkiewicz J, Paltoglou G, Buckley K, Thelwall M (2011b) Negative emotions boost user activity at BBC forum. Phys A 390(16):2936–2944

Dalgleish T, Power M (1999) Handbook of cognition and emotion. Wiley, New York

Diakopoulos NA, Shamma DA (2010) Characterizing debate performance via aggregated Twitter sentiment. In: Proceedings of CHI'10, Atlanta, GA, pp 1195–1198

Ding X, Liu B, Yu PS (2008) A holistic lexicon-based approach to opinion mining. In: Proceedings of WSDM'08, Palo Alto, CA, pp 231–240

Dodds P, Danforth C (2009) Measuring the happiness of large-scale written expression: songs, blogs, and presidents. J Happiness Stud. doi:10.1007/s10902-009-9150-0

Fan RE, Chang KW, Hsieh CJ, Wang XR, Lin CJ (2008) LIBLINEAR: a library for large linear classification. J Mach Learn Res 9:1871–1874

Fouche G (2011) Nobel peace prize may recognise Arab spring. (28 Sept, 2011, Reuters). http://in.reuters.com/article/2011/09/27/idININdia-59582020110927. Accessed 27 Dec 2011

Gazzar SE, Vitorovich L, Bender R (2011) Egypt communications cut ahead of further protests (28 Jan 2011, The Wall Street Journal). http://online.wsj.com/article/BT-CO-20110128-706943.html. Accessed 27 Dec 2011

Godbole N, Srinivasaiah M, Skiena S (2007) Large-scale sentiment analysis for news and blogs. In: Proceedings of ICWSM'07, Boulder, CO

Gonzalez-Bailon S, Paltoglou G (2012) The positive effects of negative emotions in online communities (under review)

Gonzalez-Bailon S, Banchs RE, Kaltenbrunner A (2010) Emotional reactions and the pulse of public opinion: Measuring the impact of political events on the sentiment of online discussions. CoRR. http://arxiv.org/abs/1009.4019

Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH (2009) The weka data mining software: an update. SIGKDD Explor Newsl 11:10–18

Halliday J, Garside J (2011) Rioting leads to cameron call for social media clampdown (11 Aug, 2011, The Guardian). http://www.guardian.co.uk/uk/2011/aug/11/cameron-call-social-media-clampdown. Accessed 27 Dec 2011

Harvey M (2010) Facebook ousts Google in us popularity (17 Mar 2010, The Sunday Times). http://technology.timeson-line.co.uk/tol/news/tech_and_web/the_web/article7064973.ece. Accessed 05 July 2010

Howard P (2011) The Arab spring's cascading effects (23 Feb 2011, Miller-McCune). http://www.miller-mccune.com/politics/the-cascading-effects-of-the-arab-spring-28575/. Accessed 27 Dec 2011

Jijkoun V, de Rijke M, Weerkamp W (2010) Generating focused topic-specific sentiment lexicons. In: Proceedings of ACL '08, Columbus, OH, pp 585–594

Joachims T (1999) Making large-scale SVM learning practical. In: Advances in kernel methods - support vector learning, vol 11. MIT Press, Cambridge, MA

John GH, Langley P (1995) Estimating continuous distributions in Bayesian classifiers. Stanford University, Stanford, CA, pp 338–345

Kramer AD (2010) An unobtrusive behavioral model of "gross national happiness". In: Proceedings of CHI'10, Atlanta, GA, pp 287–290

Le Cessie S, Van Houwelingen JC (1992) Ridge estimators in logistic regression. Appl Stat 41(1):191–201

Manning CD, Schütze H (1999) Foundations of statistical natural language processing. MIT Press, Cambridge, MA

Mauss IB, Robinson MD (2009) Measures of emotion: a review. Cogn Emotion 23(2):209–237

McCallum AK (2002) Mallet: a machine learning for language toolkit. http://www.cs.umass.edu/mccallum/mallet

Miller GA (1995) Wordnet: a lexical database for English. Commun ACM 38(11):39–41

Mishne G (2005) Experiments with mood classification in blog posts. In: 1st Workshop on stylistic analysis of text for information access, Salvador, Brazil

Mishne G, de Rijke M (2006) Capturing global mood levels using blog posts. In: Proceedings of AAAI-CAAW, Stanford University, Stanford, CA, pp 145–152

Mitchell TM (1997) Machine learning, 1st edn. McGraw-Hill, New York

Mitrović M, Paltoglou G, Tadić B (2011) Quantitative analysis of bloggers' collective behavior powered by emotions. J Stat Mech Theory Exp 2011(02):P02005

O'Connor B, Balasubramanyan R, Routledge BR, Smith NA (2010) From tweets to polls: linking text sentiment to public opinion time series. In: Proceedings of ICWSM'10, Washington, DC

Owsley S, Sood S, Hammond KJ (2006) Domain specific affective classification of documents. In: Proceedings of AAAICAAW'06, Stanford University, Stanford, CA, pp 181–183

Pak A, Paroubek P (2010) Twitter as a corpus for sentiment analysis and opinion mining. In: Proceedings of LREC'10, Valletta, Malta

Paltoglou G, Thelwall M (2012) Twitter, myspace, digg: unsupervised sentiment analysis in social media. ACM TIST 3(4):66.1–66.19

Paltoglou G, Thelwall M, Buckely K (2010) Online textual communication annotated with grades of emotion strength. In: Proceedings of EMOTION, Imperial College, London, pp 25–31

Paltoglou G, Theunis M, Kappas A, Thelwall M (2013) Predicting emotional responses to long informal text. J IEEE Trans Affect Comput 99(PrePrints):1

Pang B, Lee L (2008) Opinion mining and sentiment analysis. Found Trends Inform Retrieval 2(1–2):1–135

Pang B, Lee L, Vaithyanathan S (2002) Thumbs up? Sentiment classification using machine learning techniques. In: Proceedings of EMNLP'02. Association for Computational Linguistics, Philadelphia, pp 79–86

Pennebaker JW (2008) Debate 3: Mccain and Obama word usage (15 Oct 2008, WordWathcers). http://wordwatchers.wordpress.com/2008/10/15/debate-3-mccain-and-obama-word-usage/. Accessed 01 Feb 2012

Pennebaker JW, Francis ME (1999) Linguistic inquiry and word count, 1st edn. Lawrence Erlbaum, Mahwah, NJ

Pennebaker JW, Persaud R (2010) The 2010 UK election: the second debate (23 Apr 2010, WordWathcers). http://wordwatchers.wordpress.com/2010/04/23/the-2010-uk-election-the-second-debate/. Accessed 01 Feb 2012

Platt, J. C. (1999). Fast training of support vector machines using sequential minimal optimization. John C. Platt Microsoft Research 1 Microsoft Way, Redmond, WA, pp 185–208

Quirk R, Greenbaum S, Leech G, Svartvik J (1985) A comprehensive grammar of the English language. Longman, London

Russell JA (1980) A circumplex model of affect. J Person Soc Psychol 39(6):1161–1178

Sebastiani F (2002) Machine learning in automated text categorization. ACM Comput Surv 34(1):1–47

Strapparava C, Mihalcea R (2007) Semeval-2007 task 14: affective text. In: Proceedings of SemEval'07, Prague, Czech Republic, pp 70–74

Strapparava C, Mihalcea R (2008) Learning to identify emotions in text. In: Proceedings of SAC'08, Fortaleza, Ceara, Brazil, pp 1556–1560

Strapparava C, Valitutti A (2004) WordNet-Affect: an affective extension of WordNet. In: Proceedings of LREC'04 (Vol 4), Lisbon, Portugal, pp 1083–1086

Taboada M, Brooke J, Tofiloski M, Voll K, Stede M (2001) Lexicon-based methods for sentiment analysis. Comput Linguistics 37:267–307

Thelwall M (2009) Myspace comments. Online Inform Rev 33(1):58–76

Thelwall M, Buckley K, Paltoglou G, Di C, Kappas A (2010) Sentiment strength detection in short informal text. JASIST 61(12):2544–2558

Thelwall M, Buckley K, Paltoglou G (2011) Sentiment in Twitter events. J Am Soc Inf Sci Technol 62:406–418

Thomas M, Pang B, Lee L (2006) Get out the vote: determining support or opposition from congressional floor-debate transcripts. In: Proceedings of EMNLP'06. Association for Computational Linguistics, Morristown, pp 327–335

Tonkin E, Pfeiffer HD, Tourte G (2012) Twitter, information sharing and the London riots. Bull Am Soc Inf Sci Tech 38(2):49–57

Vapnik VN (1999) The nature of statistical learning theory (information science and statistics). Springer, Heidelberg

Velikovich L, Blair-Goldensohn S, Hannan K, McDonald R (2010) The viability of web-derived polarity lexicons. In: Proceedings of HLT'10, Stroudsburg, PA, USA, pp 777–785

Whitelaw C, Garg N, Argamon S (2005) Using appraisal groups for sentiment analysis. In: Proceedings of CIKM'05, Bremen, Germany, pp 625–631

Wiebe J, Bruce RF, O'Hara TP (1999) Development and use of a gold-standard data set for subjectivity classifications. In: Proceedings of 37th annual meeting of ACL, College Park, MD, USA, pp 246–253

Wilson T, Hoffmann P, Somasundaran S, Kessler J, Wiebe J, Choi Y (2005a) Opinionfinder: a system for subjectivity analysis. In: Proceedings of HLT/EMNLP-demo, Vancouver, BC, Canada, pp 34–35

Wilson T, Wiebe J, Hoffmann P (2005b) Recognizing contextual polarity in phrase-level sentiment analysis. In: Proceedings of EMNLP'05, Vancouver, BC, Canada, pp 347–354

Zaidan O, Eisner J, Piatko CD (2007) Using "annotator rationales" to improve machine learning for text categorization. In: Proceedings of HLT-NAACL. Association of Computational Linguistics, Rochester, NY, pp 260–267