Distribution of Emotional Reactions to News Articles in Twitter

Omar Juárez Gambino*[†], Hiram Calvo[†], Consuelo-Varinia García-Mendoza*

*Instituto Politécnico Nacional - ESCOM, J.D. Bátiz, Lindavista, G.A. Madero, Mexico City, 07738, Mexico
[†]Instituto Politécnico Nacional - CIC, J.D. Bátiz, Nueva Industrial Vallejo, G.A. Madero, Mexico City, 07738, Mexico
b150697@sagitario.cic.ipn.mx, hcalvo@cic.ipn.mx, cvgarcia@ipn.mx

Abstract

Several datasets of opinions expressed by Social networks' users have been created to explore Sentiment Analysis tasks like Sentiment Polarity and Emotion Mining. Most of these datasets are focused on the writers' perspective, that is, the post written by a user is analyzed to determine the expressed sentiment on it. This kind of datasets do not consider the source that provokes those opinions (e.g. a previous post). In this work, we propose a dataset focused on the readers' perspective. The developed dataset contains news articles published by three newspapers and the distribution of six predefined emotions expressed by readers of the articles in Twitter. This dataset was built aiming to explore how the six emotions are expressed by Twitter users' after reading a news article. We show some results of a machine learning method used to predict the distribution of emotions in unseen news articles.

Keywords: Twitter Sentiment Analysis, Reader's emotions, Emotion distribution

1. Introduction

Social media have fostered new ways of interaction between writers and readers. The writer of a post can receive several responses from readers as soon as it is available, and then a direct interchange of opinions begins. This particular form of communication has attracted attention of the Sentiment Analysis community because both writers and readers express emotions during their conversation. Twitter has become a popular media where users interchange opinions with strong emotional content. Since 2013 SemEval workshop has organized a Twitter Sentiment Analysis task (Nakov et al., 2013). Specific objectives of this task have changed over time, but all released datasets since 2013 have contained tweets and annotations related to sentiment polarity or emotions expressed on them. The Spanish Society of Natural Language Processing (SEPLN for its acronym in Spanish) has organized a similar task but specifically for Spanish since 2012 (Villena Román et al., 2013). The dataset for this workshop is composed by tweets and annotations corresponding to the sentiment polarity of each one. On both workshops the datasets were created focusing on writer's perspective; this means that the proposed tasks use the text written in the opinions to predict the sentiment polarity or emotions on them.

On the other hand, there have been fewer efforts focused on the reader's perspective. In this case, the idea is to determine which emotions or sentiment polarity would be expressed by a user as a reaction after reading a post. From this perspective, datasets comprise the written text read by users (e.g. a news article) and annotations corresponding to sentiment polarity or emotions expressed by the users. Predicting the sentiment polarity or emotions of readers is considered particularly difficult, because the text where readers express their opinions is not available; instead, the written text that was read by them is used for the prediction. Lin et al. (2007) and Lin et al. (2008) created datasets with news articles from Yahoo! China and the emotions users expressed after reading the articles, from a fixed set of 8 emotions. The authors pointed out that the emotions expressed in the news content are not necessarily the same as those expressed by the readers; this is an important characteristic that should be considered. Rao (2016) and Li et al. (2017) collected datasets with news articles and user ratings across 8 emotions from Sina website[1]. The tasks defined for these datasets are to predict the predominant emotion (also known as Acc@1) of each article and the percentage of votes that users will express for the defined set of emotions in each article (distribution of emotions). It is worth to mention that in all these three works, readers express their emotions by choosing emoticons from a fixed set defined by the websites. Authors argue that emoticons are related to some specific emotions, even though readers do not express their opinions in a written form.

In this work, we propose a dataset composed by news articles from three different newspapers and the corresponding distribution of emotions, from a predefined set of six emotions, expressed in texts by readers in Twitter. Our dataset is focused on reader's perspective but, unlike the aforementioned works above, the emotions are labeled by human annotators by analyzing the content of each tweet instead of a fixed emoticon set. To our knowledge there is no available dataset of news articles with their corresponding distribution of emotions from Twitter users.

2. Dataset development

News articles report information about current events. Journalists are expected to describe these events based on real facts and following three important principles: impartiality, neutrality and objectivity. From this point of view, news articles are supposed to lack personal opinions, emotions or anything additional to the reported facts. Although there are no explicit emotions in news, reported facts may trigger emotions in readers. For instance, consider the following headline news extracted from a newspaper. "Call for punishment for man who intentionally hit dog" (originally in Spanish). In the example it is clear that a factual event is reported, but a set of emotions like sadness and anger could be triggered on readers because of the news content. Therefore, we considered that news articles are a good source of emotional reactions, specially those published in Twitter.

---

[1]http://news.sina.com.cn/society/

2.1. Newspaper selection

There are several newspapers that have an online version of their articles. Some of them require a paid subscription to access articles while others have free access. We considered that free access allows more users to read the articles, so more and better responses can be available. Therefore, we discarded newspapers that required a paid subscription. Another aspect to consider is the kind of content that newspapers publish. There are newspapers specialized on subjects like finance or sports, while others include general information of different topics. We do not want to limit the scope of the dataset to specialized subjects because we think that any kind of news contents can trigger emotions on readers. The last consideration is more subjective because it is related to the bias of newspapers. We have identified that certain sensible topics like politics are published with different focus. Even though it is not our intention to affirm that a specific newspaper has bias to certain political preferences, we want to include diversity of opinions in the dataset. So at the end, we have selected three Mexican newspapers that comply with these characteristics. El Universal[2], La Jornada[3] and Excelsior[4]. In order to verify all of the above assumptions, a manual review of the collected data was carried out, finding that they were suitable for our purposes.

The three selected newspapers have active Twitter user accounts. These accounts are used to publish headlines of news articles and to share them with other Twitter users. Once users read an article, tweet responses (hereinafter replies) are posted. By analyzing these replies it is possible to identify emotions that the article triggered on users. Therefore, in our proposal we consider that it is possible to use replies as readers' emotional reactions to news articles. To illustrate this interaction between newspaper and readers, we present the following example.

News article headline published by La Jornada (originally in Spanish): "@lajornadaonline: He is not my president!, they shout in US cities after Trump's victory." A sample of its replies (translated from Spanish):

- @user1: @lajornadaonline @realDonaldTrump They are making pressure, and it is obvious that this is not necessarily going to change the state of things...

- @user2: @lajornadaonline Polarization has generated radicalizations that threaten to rise tone

- @user3: @lajornadaonline @realDonaldTrump That is democracy, accept it; it was a plot.

2.2. News and replies gathering

Twitter provides official APIs to access posts from different programming languages. The most serious limitation of the them is that there is no method available to get all post's replies, even though this request was made since 2008 to Twitter developers[5]. We have implemented an indirect pro-

cess to get news articles and their replies. The pseudocode of the process is shown in Algorithm 1.

Algorithm 1: Algorithm to get news articles and their replies

---
1 function getNewsAndReplies (q);
   Input : q is the query that specifies the data we want to get
   Output: A list of news and their replies
2 replies = searchReplies(q);
3 **foreach** *reply in replies* **do**
4   replyID = reply.ID;
5   replyContent = reply.content;
6   newsID = getInReplyTo(replyID);
7   url = getURLNews(newsID);
8   newsContent = getNewsContent(url);
9   save(newsContent, replyContent);
10 **end**

As can be seen in this algorithm, instead of obtaining the contents of the news and then the contents of the replies, we have to proceed the other way around. Contents of the news article are obtained by parsing the source code of the web page specified in the URL of the tweet. The procedure for gathering news and tweets was applied to the three selected newspapers and in the query q of Algorithm 1 we specified to recover replies published from 01-01-2016 to 01-01-2017. Table 1 shows the number of news articles and replies collected by newspaper. The number of replies may vary between newspapers depending on the number of readers the newspaper had and the impact of the articles.

Table 1: Number of news articles and replies collected

Newspaper	News articles	Replies
El Universal	90	1,000
Excelsior	100	1,136
La Jornada	98	1,406
Total	*288*	*3,542*

2.3. Set of emotions selection

The study of emotions is a multidisciplinary field which includes Neuroscience, Psychology and Cognitive Sciences. Every discipline has its own objectives and perspectives, in such a way that it has been historically hard to agree in a definition of what an emotion is (Kleinginna and Kleinginna, 1981). Nevertheless, there are well-founded proposals on types of emotions and how they are generated (Ortony and Turner, 1990). A remarkable study of emotions was done in Shaver et al. (1987), in which the authors argue that emotional knowledge should be represented hierarchically. To illustrate that, an experiment was developed, in which 112 Psychology students were asked to rate 213 words used to express emotions. By applying mean prototypicality rating, 135 words were selected as good emotions descriptors. Then another 100 psychology students were asked to do a similarity sorting over the

135 words. The sorting resulted in a hierarchy with six main groups and every group had 2 internal groups. At the end, the authors proposed a 3 level hierarchy, with the first level containing basic emotions, while the second and third levels, represented more specific emotions. In Tables 2 and 3 we show the emotions, translated to Spanish, of every group and its corresponding level. It is important to mention that the number of words used to describe emotions was reduced from the original 135 to 110 because some of them are translated using the same word in Spanish [6]. We will adhere to this proposal and the words listed in the tables will be used as labels to tag the emotions in replies during the annotation process.

2.4. Annotation process

In order to identify emotions in replies, we selected 4 annotators: 3 undergraduate Computer Science students and a Computer Science professor. They were provided with 288 news articles collected from the three selected newspapers and 3,542 replies (with an average of 11 replies per article). Annotators were asked to follow the next procedure for each news article:

1. Read each reply to the article.

2. Identify the emotions expressed in each reply.

3. Mark all emotions identified in the replies using the words (tags) of Tables 2 and 3.

It was emphasized to the annotators that the emotions that should be annotated are those expressed in the replies, and not those that the news article triggered in them.

According to their experience when tagging emotions, it was easier for annotators to have an extensive list of labels than only the six ones defined in level 1 of each group: *love*, *joy*, *surprise*, *anger*, *sadness* and *fear*. As a result, we obtained a fine-grained annotated dataset. Although identifying very specific emotions such as *loathing* and *exasperation* is interesting, in this work we want to predict the general emotions expressed by Twitter users after reading a news article[7]. Therefore, all the emotions marked by the annotators were generalized to the emotions in level 1 because they are considered basic emotions and provide a good overall idea of the sentiment expressed in replies. To generalize emotions, we used the defined groups and replaced the words of levels 2 and 3 with the word in level 1. For instance, let us consider that an annotator tagged a reply with the labels *enjoyment*, *relief* and *surprise*. By looking in group 2, we identified that the labels enjoyment and relief are related to the basic emotion *joy*, while *surprise* is already the basic emotion of group 3 and hence the generalized tags are *joy* and *surprise*. These generalized tags are the ones that were used in the created dataset.

It is important to measure the agreement in tasks performed by humans. The higher the agreement, the better the quality of the dataset; but in this particular task there are some factors that make difficult to expect a high agreement. First of

Table 2: Emotions provided to annotators

Group 1	
Level	Emotions
1	**Amor (Love)**
2	Afecto (Affection), Deseo (Lust), Anhelo (Longing)
3	Adoración (Adoration), Amor (Love), Cariño (Fondness), Afición (Liking), Atracción (Attraction), Ternura (Tenderness), Compasión (Compassion), Sentimentalismo (Sentimentality), Agitación (Arousal), Deseo (Lust), Pasión (Passion), Encaprichamiento (Infatuation), Anhelo (Longing)
Group 2	
Level	Emotions
1	**Alegría (Joy)**
2	Alegría (Joy), Entusiasmo (Zest), Contentamiento (Contentment), Orgullo (Pride), Optimismo (Optimism), Encanto (Enthrallment), Alivio (Relief)
3	Diversión (Amusement), Dicha (Bliss), Alegría (Cheerfulness), Regocijo (Glee), Jovialidad (Joviality), Deleite (Delight), Placer (Enjoyment), Felicidad (Happiness), Júbilo (Jubilation), Euforia (Elation), Satisfacción (Satisfaction), Éxtasis (Ecstasy)
Group 3	
Level	Emotions
1	**Sorpresa (Surprise)**
2	Sorpresa (Surprise)
3	Asombro (Amazement), Sorpresa (Surprise), Estupor (Astonishment)
Group 4	
Level	Emotions
1	**Enojo (Anger)**
2	Irritación (Irritation), Exasperación (Exasperation), Rabia (Rage), Disgusto (Disgust), Envidia (Envy), Tormento (Torment)
3	Irritación (Irritation), Agitación (Agitation), Molestia (Annoyance), Mal humor (Grouchiness), Exasperación (Exasperation), Enojo (Anger), Rabia (Rage), Furia (Fury), Ira (Wrath), Hostilidad (Hostility), Agresividad (Ferocity), Rencor (Bitterness), Odio (Hate), Aversión (Loathing), Desprecio (Scorn), Desagrado (Dislike), Resentimiento (Resentment), Disgusto (Disgust), Envidia (Envy), Celos (Jealousy), Tormento (Torment)

all, the content of news articles often leads to controversial opinions, for instance political stances, and in these kind of opinions is not easy to identify emotions because they are not explicit expressed or they contain sarcasm. In addition to that, the background of annotators may lead to variation in the identified emotions. Finally, because of the number of annotators and emotions to identify, a perfect agreement

---

[6] For further details of the original hierarchy and emotions see (Shaver et al., 1987)

[7] Prediction of emotions in fine-grained level as well as the use of the hierarchy are considered as future work

Table 3: Emotions provided to annotators 2

Group 5	
Level	Emotions
1	**Tristeza (Sadness)**
2	Sufrimiento (Suffering), Tristeza (Sadness), Decepción (Disappoinment), Pena (Shame), Desamparo (Neglect), Lástima (Sympathy)
3	Agonía (Agony), Sufrimiento (Suffering), Dolor (Hurt), Angustia (Anguish), Depresión (Depression), Desesperación (Despair), Desesperanza (Hopelessness), Pesadumbre (Gloom), Abatimiento (Glumness), Tristeza (Sadness), Infelicidad (Unhappiness), Aflicción (Grief), Pesar (Sorrow), Miseria (Misery), Melancolía (Melancholy), Consternación (Dismay), Decepción (Disappointment), Disgusto (Displeasure), Culpa (Guilt), Pena (Shame), Arrepentimiento (Regret), Remordimiento (Remorse), Aislamiento (Alienation), Desamparo (Neglect), Soledad (Loneliness), Rechazo (Rejection), Nostalgia (Homesickness), Derrota (Defeat), Abatimiento (Dejection), Inseguridad (Insecurity), Vergüenza (Embarrassment), Humillación (Humiliation), Insulto (Insult), Pena (Pity), Lástima (Sympathy)
Group 6	
Level	Emotions
1	**Miedo (Fear)**
2	Horror (Horror), Ansiedad (Nervousness)
3	Sobresalto (Alarm), Conmoción (Shock), Miedo (Fear), Temor (Fright), Horror (Horror), Terror (Terror), Pánico (Panic), Histeria (Hysteria), Mortificación (Mortification), Ansiedad (Anxiety), Nerviosismo (Nervousness), Tensión (Tenseness), Inquietud (Uneasiness), Aprensión (Apprehension), Preocupación (Worry), Angustia (Distress)

happens when the four annotators tag a reply with the same emotions from a set of six non-exclusive labels (love, joy, surprise, anger, sadness and fear) and this is already an unlikely condition. Despite all the above, the inter-annotator agreement was 0.48 using a generalized version of Cohen's kappa (multi-kappa) (Davies and Fleiss, 1982). This value is considered by Landis and Koch (1977) as a moderated inter-annotator agreement and we also consider it so for the purposes of this work.

2.5. Distribution of emotions

Users that reply to news articles can express multiple emotions in their posts. These emotions can be counted in order to determine the frequency (*votes*) of each emotion. To illustrate this idea, let us consider that a news article had ten replies and for each reply a set of the six emotions previously defined have been tagged by the annotators. Table 4 shows tagged emotions indicated with a checkmark.

It is possible to represent the information of Table 4 in terms of percentage applying a normalization process, dividing each count of tagged emotions by the number of replies, as shown in the last row. These percentage values is what we call *distribution of emotional reaction* and can be interpreted as the degree on which of these emotions were expressed by readers as a reaction to a news article. For instance, the news article of Table 4 provoked 50% of love, 100% of joy, 40% of surprise, 30% of sadness, 70% of anger and 0% of fear. It is worth mentioning that percentage values of emotions are independent because a news article may provoke 100% of joy and 70% of anger at the same time. Therefore, the sum of these values do not necessarily sum 100%. We have decided to use percentage ranges instead of total votes to specify the distribution of emotions in the dataset. The possible values for each emotion is composed by a set of 11 ranges from 0% to 100% with 10% increment, that is {0%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%}. The same normalization process was applied to each set of responses in order to have all news articles associated with their corresponding distribution of emotions.

In Figures 1, 2, 3 we show the distribution of percentage values per emotion in articles of each newspaper. We have also created an extra dataset that integrates the news articles and their corresponding tags of the three newspaper, which allows to test a more generalized prediction of emotion distribution. Emotions have different distribution values in news articles as can be seen in Figure 1, where 50% of anger was expressed as a reaction in 7 articles while the same percentage of sadness was expressed in 20 articles. In addition to that, the integrated version of newspapers (see Figure 4) clearly shows that some emotions are less frequently expressed in replies to news articles. Particularly love and surprise were not expressed at all (0%) in 229 and 189 articles respectively, it means that less than 40% of articles provoked those emotions. This characteristic leads to a unbalanced corpus in which some percentage values are much more frequent than others.
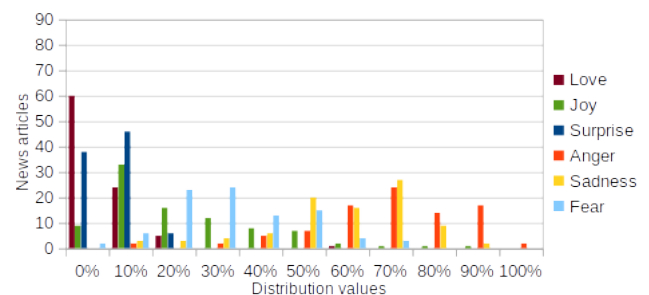

Figure 1: El Universal.

3. Data format

Data are presented in four different CSV text files. Each file is related to a specific newspaper and the integrated version of the three newspapers. Files contain one news article per line. Each line has three separated elements: tweet ID, the content of the news article, and the tagged distribution of emotions. Elements are separated by the

Table 4: Reactions of users to a news article

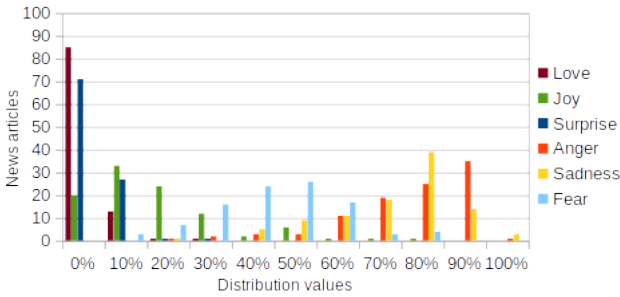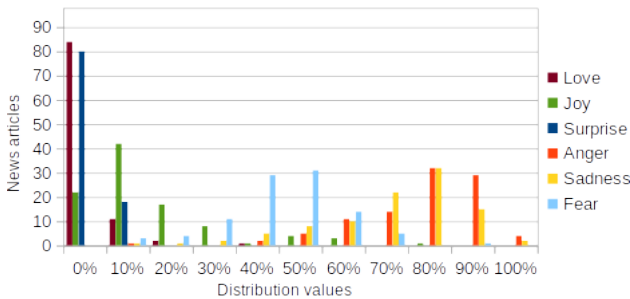| Responses	Emotions					
	Love	Joy	Surprise	Sadness	Anger	Fear
R_1		✓	✓			
R_2	✓	✓		✓	✓	
R_3		✓			✓	
R_4	✓	✓	✓		✓	
R_5		✓				
R_6	✓	✓			✓	
R_7		✓	✓	✓		
R_8	✓	✓			✓	
R_9		✓			✓	
R_{10}	✓	✓	✓	✓	✓	
Total	*5/10*	*10/10*	*4/10*	*3/10*	*7/10*	*0/10*
Normalized total	*0.5*	*1*	*0.4*	*0.3*	*0.7*	*0*

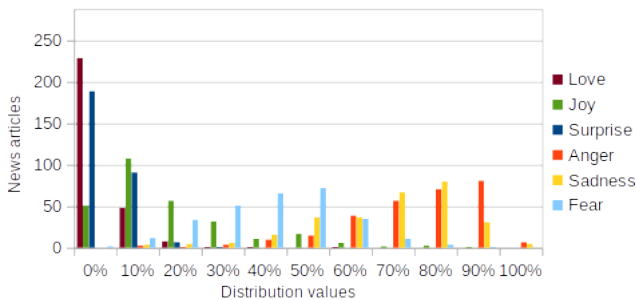Figure 2: Excelsior.

Figure 3: La Jornada.

Figure 4: Integrated newspapers.

string "&&". A tuple of six elements is used to describe the distribution of emotions. The order of the elements in the tuples is associated with the emotions love, joy, surprise, anger, sadness and fear respectively. The values of the distribution of emotions are expressed in decimal form. An example of an annotated news article is the following: 123&&This is the content of the article&&0.1,0.2,0.1,0.7,0.8,0.5

4. Possible uses and early experiments

The distribution of emotions annotated in the proposed dataset is useful for several Sentiment Analysis tasks. Let us consider the distribution of emotions shown in Table 4 (50%, 100%, 40%, 30%, 70% and 0%). Given a news article, it is possible to determine the predominant emotion (Joy)—cf. (Lin et al., 2007), determine the ranking of emotions (Joy, Anger, Love, Surprise, Sadness and Fear)—cf. (Lin et al., 2008) and of course the very same distribution of emotions —cf. (Rao, 2016), (Li et al., 2017). Practical applications of predicting the distribution of emotions could be the following:

- Assisted writing for provoking particular emotions.

- Recover news articles that generate specific emotions (e.g. news articles that cause surprise)

In order to explore how machine learning methods could use this dataset for predicting distribution of emotions, we have performed some experiments. Because users can express more than one emotion in their replies and each emotion is associated with 11 possible distribution values (from 0% to 100%), we decided to use a multi-target classification strategy (Herrera et al., 2016) that supports multi-dimensional problems such as this. This strategy follows a supervised approach by using well known methods like Naïve Bayes and SVM as base classifiers. As a metric we used the average Pearson Correlation (AP). This metric has been used in (Rao, 2016) and (Li et al., 2017) to compare the predicted distribution of emotions against the real ones. Values of this metric ranges from -1 to 1, where 1 indicates a perfect positive correlation. Table 5 shows the average results from 10-fold cross validation, using 90% of data for training and 10% for testing. We believe that, when a newspaper has a clear political bias (left-wing or right-wing),

most of its readers will share the same ideologies and they will express similar emotions in their replies. On the other hand, when a newspaper has an intermediate position (centrist), the opinions expressed by its readers will be diverse as well as their emotions. This idea is reflected in the results of Table 5 in which La Jornada (left-wing) obtained the best results, following by Excelsior (right-wing), and finally El Universal (centrist). The Integrated version is a kind of average between the results of the three newspapers. Further details of the experiments are described in (Gambino and Calvo, 2018).

Table 5: AP results by each newspaper

Newspaper	AP
Excelsior	0.9241
La Jornada	0.9274
El Universal	0.8494
Integrated	0.8997

5. Conclusions and future work

In this work we presented a dataset of news articles and their corresponding distribution of emotions expressed by readers in Twitter. This dataset could be used for several Sentiment Analysis tasks. The expressed emotions are from the readers' perspective; and, the available datasets of this kind are scarce. Early experiments were performed to explore the application of machine learning methods for predicting emotion distribution and results seem promising. Increasing the size of the corpus is proposed as a future work.

6. Acknowledgements

7. Bibliographical References

Davies, M. and Fleiss, J. L. (1982). Measuring Agreement for Multinomial Data. *Biometrics*, 38(4):1047–1051.

Gambino, O. J. and Calvo, H. (2018). Modeling distribution of emotional reactions in social media using a multi-target strategy. *Jounal of Intelligent & Fuzzy Systems*, in press.

Herrera, F., Charte, F., Rivera, A. J., and Del Jesus, M. J. (2016). *Multilabel Classification: Problem Analysis, Metrics and Techniques*. Springer International Publishing.

Kleinginna, P. R. and Kleinginna, A. M. (1981). A Categorized List of Emotion Definitions, with Suggestions for a Consensual Definition. *Motivation and emotion*, 5(4):345–379.

Landis, J. R. and Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33(1):159–174.

Li, X., Peng, Q., Sun, Z., Chai, L., and Wang, Y. (2017). Predicting Social Emotions from Readers' Perspective. *IEEE Transactions on Affective Computing*, in press.

Lin, K. H.-Y., Yang, C., and Chen, H.-H. (2007). What emotions do news articles trigger in their readers? In *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 733–734. ACM.

Lin, K. H.-Y., Yang, C., and Chen, H.-H. (2008). Emotion classification of online news articles from the reader's perspective. In *International Conference on Web Intelligence and Intelligent Agent Technology, 2008.* , volume 1, pages 220–226. IEEE.

Nakov, P., Rosenthal, S., Kozareva, Z., Stoyanov, V., Ritter, A., and Wilson, T. (2013). Semeval-2013 task 2: Sentiment analysis in twitter. In *Proceedings of the Seventh International Workshop on Semantic Evaluation (SemEval 2013)*, volume 2, pages 312–320.

Ortony, A. and Turner, T. J. (1990). What's basic about basic emotions? *Psychological review*, 97(3):315.

Rao, Y. (2016). Contextual sentiment topic model for adaptive social emotion classification. *IEEE Intelligent Systems*, 31(1):41–47.

Shaver, P., Schwartz, J., Kirson, D., and O'connor, C. (1987). Emotion Knowledge: Further Exploration of a Prototype Approach. *Journal of personality and social psychology*, 52(6):1061.

Villena Román, J., Lana Serrano, S., Martínez Cámara, E., and González Cristóbal, J. C. (2013). TASS - Workshop on Sentiment Analysis at SEPLN. *Procesamiento del Lenguaje Natural*, (50):37–44.