Multimodal Sentiment Analysis of Social Media

Diana Maynard[1], David Dupplaw[2], and Jonathon Hare[2]

[1] University of Sheffield, Department of Computer Science
Regent Court, 211 Portobello, Sheffield, S1 4DP, UK
diana@dcs.shef.ac.uk
[2] University of Southampton, Web and Internet Science
Southampton, Hampshire, SO17 1BJ, UK
dpd|jsh2@ecs.soton.ac.uk

Abstract This paper describes the approach we take to the analysis of social media, combining opinion mining from text and multimedia (images, videos, etc), and centred on entity and event recognition. We examine a particular use case, which is to help archivists select material for inclusion in an archive of social media for preserving community memories, moving towards structured preservation around semantic categories. The textual approach we take is rule-based and builds on a number of sub-components, taking into account issues inherent in social media such as noisy ungrammatical text, use of swear words, sarcasm etc. The analysis of multimedia content complements this work in order to help resolve ambiguity and to provide further contextual information. We provide two main innovations in this work: first, the novel combination of text and multimedia opinion mining tools; and second, the adaptation of NLP tools for opinion mining specific to the problems of social media.

1 Introduction

Social web analysis is all about the users who are actively engaged and generate content. This content is dynamic, reflecting the societal and sentimental fluctuations of the authors as well as the ever-changing use of language. Social networks are pools of a wide range of articulation methods, from simple "Like" buttons to complete articles, their content representing the diversity of opinions of the public. User activities on social networking sites are often triggered by specific events and related entities (e.g. sports events, celebrations, crises, news articles) and topics (e.g. global warming, financial crisis, swine flu).

With the rapidly growing volume of resources on the Web, archiving this material becomes an important challenge. The notion of community memories extends traditional Web archives with related data from a variety of sources. In order to include this information, a semantically-aware and socially-driven preservation model is a natural way to go: the exploitation of Web 2.0 and the wisdom of crowds can make web archiving a more selective and meaning-based process. The analysis of social media can help archivists select material for inclusion, while social media mining can enrich archives, moving towards structured preservation around semantic categories. In this paper, we focus on the challenges in the development of opinion mining tools from both textual and multimedia content.

We focus on two very different domains: socially aware federated political archiving (realised by the national parliaments of Greece and Austria), and socially contextualized broadcaster web archiving (realised by two large multimedia broadcasting organizations based in Germany: Sudwestrundfunk and Deutsche Welle). The aim is to help journalists and archivists answer questions such as what the opinions are on crucial social events, how they are distributed, how they have evolved, who the opinion leaders are, and what their impact and influence is.

Alongside natural language, a large number of the interactions which occur between social web participants include other media, in particular images. Determining whether a specific non-textual media item is performing as an opinion-forming device in some interaction becomes an important challenge, more so when the textual content of some interaction is small or has no strong sentiment. Attempting to determine a sentiment value for an image clearly presents great challenges, and this field of research is still in its infancy. We describe here some work we have been undertaking, firstly to attempt to provide a sentiment value from an image outside of any specific context, and secondly to utilise the multimodal nature of the social web to assist the sentiment analysis of either the multimedia or the text.

2 Related Work

While much work has recently focused on the analysis of social media in order to get a feel for what people think about current topics of interest, there are, however, still many challenges to be faced. State of the art opinion mining approaches that focus on product reviews and so on are not necessarily suitable for our task, partly because they typically operate within a single narrow domain, and partly because the target of the opinion is either known in advance or at least has a limited subset (e.g. film titles, product names, companies, political parties, etc.).

In general, sentiment detection techniques can be roughly divided into lexicon-based methods [22] and machine-learning methods, e.g. [1]. Lexicon-based methods rely on a sentiment lexicon, a collection of known and pre-compiled sentiment terms. Machine learning approaches make use of syntactic and/or linguistic features, and hybrid approaches are very common, with sentiment lexicons playing a key role in the majority of methods. For example, [17] establish the polarity of reviews by identifying the polarity of the adjectives that appear in them, with a reported accuracy of about 10% higher than pure machine learning techniques. However, such relatively successful techniques often fail when moved to new domains or text types, because they are inflexible regarding the ambiguity of sentiment terms. The context in which a term is used can change its meaning, particularly for adjectives in sentiment lexicons [18]. Several evaluations have shown the usefulness of contextual information [26], and have identified context words with a high impact on the polarity of ambiguous terms [8]. A further bottleneck is the time-consuming creation of these sentiment dictionaries, though solutions have been proposed in the form of crowdsourcing techniques[3].

Almost all the work on opinion mining from Twitter has used machine learning techniques. [19] aimed to classify arbitrary tweets on the basis of positive, negative and neutral sentiment, constructing a simple binary classifier which used n-gram and POS features, and trained on instances which had been annotated according to the existence of positive and negative emoticons. Their approach has much in common with an earlier sentiment classifier constructed by [9], which also used unigrams, bigrams and POS tags, though the former demonstrated through analysis that the distribution of certain POS tags varies between positive and negative posts. One of the reasons for the relative paucity of linguistic techniques for opinion mining on social media is most likely due to the difficulties in using NLP on low quality text [7]; for example. the Stanford NER drops from 90.8% F1 to 45.88% when applied to a corpus of tweets [14].

There have been a number of recent works attempting to detect sarcasm in tweets and other user-generated content [23, 13, 20, 5], with accuracy typically around 70-80%. These mostly train over a set of tweets with the #sarcasm and/or #irony hashtags, but all simply try to classify whether a sentence or tweet is sarcastic or not (and occasionally, into a set of pre-defined sarcasm types). However, none of these approaches go beyond the initial classification step and thus cannot predict how the sarcasm will affect the sentiment expressed. This is one of the issues that we tackle in our work.

Extracting sentiment from images is still a research area that is in its infancy and not yet prolifically published. However, those published often use small datasets for their ground truth on which to build SVM classifiers. Evaluations show systems often respond only a little better than chance for trained emotions from general images [27]. The implication is that the feature selection for such classification is difficult. [25] used a set of colour features for classifying their small ground-truth dataset, also using SVMs, and publish an accuracy of around 87%. In our work, we expand this colour-based approach to use other features and also use the wisdom of the crowd for selecting a large ground-truth dataset.

Other papers have begun to hint at the multimodal nature of web-based image sentiment. Earlier work, such as [11], is concerned with similar multimodal image annotation, but not specifically for sentiment. They use latent semantic spaces for correlating image features and text in a single feature space. In this paper, we describe the work we have been undertaking in using text and images together to form sentiment for social media.

3 Opinion Mining from Text

3.1 Challenges

There are many challenges inherent in applying typical opinion mining and sentiment analysis techniques to social media. Microposts such as tweets are, in

[3] `http://apps.facebook.com/sentiment-quiz`

some sense, the most challenging text type for text mining tools, and in particular for opinion mining, since the genre is noisy, documents have little context and assume much implicit knowledge, and utterances are often short. As such, conventional NLP tools typically do not perform well when faced with tweets [2], and their performance also negatively affects any following processing steps.

Ambiguity is a particular problem for tweets, since we cannot easily make use of coreference information: unlike in blog posts and comments, tweets do not typically follow a conversation thread, and appear much more in isolation from other tweets. They also exhibit much more language variation, and make frequent use of emoticons, abbreviations and hashtags, which can form an important part of the meaning. Typically, they also contain extensive use of irony and sarcasm, which are particularly difficult for a machine to detect. On the other hand, their terseness can also be beneficial in focusing the topics more explicitly: it is very rare for a single tweet to be related to more than one topic, which can thus aid disambiguation by emphasising situational relatedness.

In longer posts such as blogs, comments on news articles and so on, a further challenge is raised by the tracking of changing and conflicting interpretations in discussion threads. We investigate first steps towards a consistent model allowing for the pinpointing of opinion holders and targets within a thread (leveraging the information on relevant entities extracted).

We refer the reader to [2] for our work on twitter-specific IE, which we use as pre-processing for the opinion mining described below. It is not just tweets that are problematic, however; sarcasm and noisy language from other social media forms also have an impact. In the following section, we demonstrate some ways in which we deal with this.

3.2 Opinion Mining Application

Our approach is a rule-based one similar to that used by [22], focusing on building up a number of sub-components which all have an effect on the score and polarity of a sentiment. In contrast, however, our opinion mining component finds opinions relating to previously identified entities and events in the text. The core opinion mining component is described in [15], so we shall only give an overview here, and focus on some issues specific to social media which were not dealt with in that work, such as sarcasm detection and hashtag decomposition.

The detection of the actual opinion is performed via a number of different phases: detecting positive, negative and neutral words, identifying factual or opinionated versus questions or doubtful statements, identifying negatives, sarcasm and irony, analysing hashtags, and detecting extra-linguistic clues such as smileys. The application involves a set of grammars which create annotations on segments of text. The grammar rules use information from gazetteers combined with linguistic features (POS tags etc.) and contextual information to build up a set of annotations and features, which can be modified at any time by further rules. The set of gazetteer lists contains useful clues and context words: for example, we have developed a gazetteer of affect/emotion words from Word-

Net [16]. The lists have been modified and extended manually to improve their quality.

Once sentiment words have been matched, we find a linguistic relation between these and an entity or event in the sentence or phrase. A Sentiment annotation is created for that entity or event, with features denoting the polarity (positive or negative) and the polarity score. Scores are based on the initial sentiment word score, and intensified or decreased by any modifiers such as swear words, adverbs, negation, sarcasm etc, as explained next.

Swear words are particularly prolific on Twitter, especially on topics such as popular culture, politics and religion, where people tend to have very strong views. To deal with these, we match against a gazetteer list of swear words and phrases, which was created manually from various lists found on the web and from manual inspection of the data, including some words acquired by collecting tweets with swear words as hashtags (which also often contain more swear words in the main text of the tweet).

Much useful sentiment information is contained within hashtags, but this is problematic to identify because hashtags typically contain multiple words within a single token, e.g. #notreally. If a hashtag is camelcased, we use the capitalisation information to create separate tokens. Second, if the hashtag is all lowercase or all uppercase, we try to form a token match against the Linux dictionary. Working from left to right, we look for the longest match against a known word, and then continue from the next offset. If a combination of matches can be found without a break, the individual components are converted to tokens. In our example, #notreally would be correctly identified as "not" + "really". However, some hashtags are ambiguous: for example, "#greatstart" gets split wrongly into the two tokens "greats" + "tart". These problems are hard to deal with; in some cases, we could make use of contextual information to assist.

We conducted an experiment to measure the accuracy of hashtag decomposition, using a corpus of 1000 tweets randomly selected from the US elections crawl that we undertook in the project. 944 hashtags were detected in this corpus, of which 408 were identified as multiword hashtags (we included combinations of letters and numbers as multiword, but not abbreviations). 281 were camelcased and/or combinations of letters and nubers, 27 were foreign words, and the remaining 100 had no obvious token-distinguishing features. Evaluation on the hard-to-recognise cases (non-camel-cased multiword hashtags) produced scores of 86.91% Precision, 90% Recall, and an F-measure of 88.43%. Given that these hard-to-resolve combinations form roughly a quarter of the multiword hashtags in our corpus, and that we are entirely successful in decomposing the remaining hashtags, this means that the overall accuracy for hashtag decomposition is much higher.

In addition to using the sentiment information from these hashtags, we also collect new hashtags that typically indicate sarcasm, since often more than one sarcastic hashtag is used. For this, we used the GATE gazetteer list collector to collect pairs of hashtags where one was known to be sarcastic, and examined the second hashtag manually. From this we were able to identify a further set of sarcasm-indicating hashtags, such as #thanksdude, #yay etc. Further investigation needs to be performed on these to check how frequently they actually indicate sarcasm when used on their own.

Finally, emoticons are processed like other sentiment-bearing words, according to another gazetteer list, if they occur in combination with an entity or event. For example, the tweet "They all voted Tory :-(" would be annotated as negative with respect to the target "Tory". Otherwise, as for swear words, if a sentence contains a smiley but no other entity or event, the sentence gets annotated as sentiment-bearing, with the value of that of the smiley from the gazetteer list.

Once all the subcomponents have been run over the text, a final output is produced for each sentiment-bearing segment, with a polarity (positive or negative) and a score, based on combining the individual scores from the various components (for example, the negation component typically reverses the polarity, the adverbial component increases the strength of the sentiment, and so on. Aggregation of sentiment then takes place for all mentions of the same entity/event in a document, so that summaries can be created.

4 Extracting Opinions from Images

4.1 Challenges

The main challenge with annotating non-textual media is that the underlying tokens within it are considerably less explicit than in textual media. In images and video, these underlying tokens are groups of pixels (compared with groups of characters [words] in text). As well as having multiple dimensions, the tokens have considerably more variation when representing exactly the same concept, and so using dictionaries and other traditional text-based tools becomes impossible. And so, we enter the world of image understanding and computer vision which, although over 30 years old, has made fewer revolutionary leaps than NLP. State of the art computer vision is still relatively basic for most general applications. This "semantic gap" between what computer vision can achieve and the level of understanding required for sentiment analysis is why extracting opinions from images is so difficult.

That said, certain specific applications have made advances recently - one of which is the application of computer vision for detecting and recognising faces of people. [24] developed a technique for face detection that is probably the most widespread computer-vision technique of all time, as most point-and-shoot cameras include face detection based on this algorithm. It uses some 1-dimensional peak features (Haar features) that are used to train a cascade of classifiers for general object detection. Trained on faces, these can detect faces in images robustly and efficiently.

Detecting the presence of a face is just the first part; fitting a model to a face can then provide some extra information about the shape and the expression of the face. Active Shape Models [3] (ASM) and Active Appearance Models [4] (AAM) are well-known algorithms for fitting a shape to an image using the image's gradients to choose the best position for the vertices of the shape. As these models are parametric and generative (they are reconstructed using a small number of parameters), a large range of poses, expressions and appearances (skin textures) can be generated. Fitting a model to an image is a constrained optimisation problem in which the parameters of the model are iteratively updated in order to minimise the difference between the generated model and the image (hence Constrained Local Model [CLM]). Once a model is fitted to an image, the parameters can then be used as input to an expression classifier that can determine an expression label for the face.

This model fits well with the Facial Action Coding System (FACS) which aims to provide a standardised way of describing the expressions of faces. Codes represent muscular actions in the face (such as "inner eyebrow raising", or "lip corner puller") and when combined they represent emotions (for example, activation of the lip corner puller AU6 and the cheek raiser AU12 actions imply happiness). These muscular movements map to combinations of parameters in the face model, so a classifier can be trained to recognise these actions. Of course, this relies on accurate face model fitting, but it is difficult to build a shape model (ASM, AAM or CLM) that will accurately fit all faces and poses, which is essential for the accurate measurement of the shape parameters needed for expression classification. Another problem is that accurate detection of a face is required to initialise the fitting of a face model; whilst face detection techniques are quite mature, they can still have major problems working in real-world images where the faces are not exactly frontal to the camera, or there are shadows or contrast issues.

4.2 Detecting Sentiment in Images

Figure 1 shows an example of a programme that recognises the expressions in a laboratory setting. In the wild, we found that inaccuracies in the face model alignment would regularly cause misclassification of the action units, and therefore the expressions.

In less constrained multimedia, we cannot rely on there being faces in the images, and sentiment may be carried by other visual traits. Indeed, images may intrinsically have sentiment associated with them through design (say a poster for a horror film) or through association with a specific subject matter which may be context sensitive (say a photo of wind generators in the context of climate change). For these situations there are no specific algorithms we can use for extracting the sentiment. However, we can perform standard feature-label correlations using classifiers over ground-truth datasets. Unfortunately, large, well labelled datasets for image sentiment are very thin on the ground. For that reason, we turned to the "wisdom of the crowd" for generating a large dataset to use for classification. Using SentiWordNet, we were able to query Flickr for the words that had the strongest positive and negative sentiments, and retrieve sets of images for each of them. Combined, these formed our ground-truth for positive and negative sentiment images. The details of this work are described in [21], but we will summarise the conclusions here.

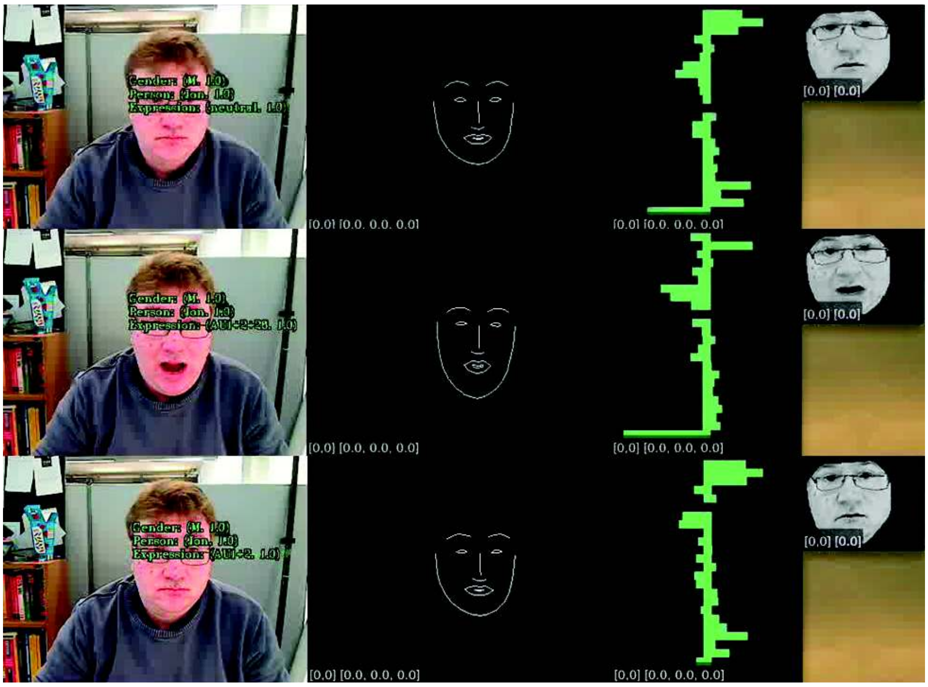

Figure 1. Recognition of expressions in a laboratory setting

We gathered images for the 1000 strongest sentiment words from SentiWordNet. This resulted in 586,000 images, most of which had a resolution of more than 1 megapixel. We extracted global and local colour features and SIFT local features from the images, and trained an SVM classifier to recognise positive/negative sentiment. We can observe that for small recall values, precision values of up to 70% can be reached. Due to the challenging character of this task, for high recall values, the precision degrades down to the random baseline. Interestingly, using mutual information, we were able to reverse engineer the correlations in the classifier to determine which features were correlated to which labels. We found that positive images had overall warm colours (reds, oranges, yellows, skin tones) and negative images had colder colours (blues, dark greens). The location of the colour had no real significance. The negative SIFT features seem dominated by a very light central blob surrounded by a much darker background, while the positive SIFT features are dominated by a dark blob on the side of the patch.

Clearly, from a basis where there is no context, it is only possible to achieve a limited amount of understanding. However, using the contextual information (e.g. co-located text) it is possible to aggregate various forms of analysis and make further estimates of an object's sentiment. To do that, it is necessary to find the text and images which are co-located. In web pages, we can extract the 'important' part of the page using boilerplate removal tools, such as our tool Readability4J [12]. In tweets, images are usually presented as links, usually to a URL shortener. It is necessary to follow the links to their destination, then to parse the final destination for a the "co-located" image. Once we have images related to the text, we look for entities within the visual content. As described in Section 3, we extract entities from the text and associate a sentiment value with them based on the textual context. These entities will be people, locations, or organisations and can be used to guide our analysis of the associated images. It is impractical to consider an entity recognition system that would recognise any entity (e.g. any person or any place), so we can use the entities in the text to reduce the search space. For example, we can use the detected person entities to train a face recognition system (for previously unseen people, on-the-fly using the image search results from major search engines), the location entites to fix a prior on a world-based search algorithm (as our work in [6]), or the organisation entities to train a logo detector.

One of the interesting insights into the social web is to see how media is spread – how it is reused and talked about and whether the sentiment associated with the media changes. We developed a system called Twitter's Visual Pulse [10] which finds near-duplicate images from a live or static Twitter stream. We used a technique called Locality Sensitive Hashing (LSH) of SIFT features extracted from the images, and determine near-duplicates by finding connected components in a graph where nodes are hashed features and edges are weighted based on the number of matching hashes. By extracting the sentiment from the tweets associated with these duplicate images, we can find out how the image is used in different contexts. In many cases, the image may be reused in contexts which are, overall, sentimentally ambivalent; however, there may be cases where an image is used in a consistent way - for example, a particular image may be used in consistently positive tweets. We form a discrete probability distribution for images falling in specific sentiment categories, which we can use to assign sentiment probabilities to the image when it is further reused, particularly in cases where the textual sentiment analysis is inconclusive.

5 Conclusions

In this paper, we have described the general approach we undertake to the analysis of social media, using a combination of textual and multimedia opinion mining tools. It is clear that both opinion mining in general, and the wider analysis of social media, are difficult tasks from both perspectives, and there are many unresolved issues. The modular nature of our approach also lends itself to new advances in a range of subtasks: from the difficulties of analysing the noisy forms of language inherent in tweets, to the problems of dealing with sarcasm in social media, to the ambiguities inherent in such forms of web content that inhibit both textual and multimedia analysis tools. Furthermore, to our knowledge this is the first system that attempts to combine such kinds of textual and multimedia analysis tools in an integrated system, and preliminary results are very promising, though this is nevertheless very much ongoing research. Future work includes further development of the opinion mining tools: we have already begun investigations into issues such as sarcasm detection, more intricate use of discourse analysis and so on.

Acknowledgments

This work was supported by the European Union under grant agreements No. 270239 Arcomem[4] and No. 610829 DecarboNet[5].

[4] http://www.arcomem.eu
[5] http://www.decarbonet.eu

Bibliography

[1] Boiy, E., Moens, M.F.: A machine learning approach to sentiment analysis in multilingual web texts. Information Retrieval 12(5), 526–558 (2009)

[2] Bontcheva, K., Derczynski, L., Funk, A., Greenwood, M.A., Maynard, D., Aswani, N.: TwitIE: An Open-Source Information Extraction Pipeline for Microblog Text. In: Proceedings of the International Conference on Recent Advances in Natural Language Processing. Association for Computational Linguistics (2013)

[3] Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J.: Active shape models – their training and application. Comput. Vis. Image Underst. 61(1), 38–59 (Jan 1995), `http://dx.doi.org/10.1006/cviu.1995.1004`

[4] Cootes, T., Edwards, G., Taylor, C.: Active appearance models. Pattern Analysis and Machine Intelligence, IEEE Transactions on 23(6), 681–685 (2001)

[5] Davidov, D., Tsur, O., Rappoport, A.: Semi-supervised recognition of sarcastic sentences in Twitter and Amazon. In: Proceedings of the Fourteenth Conference on Computational Natural Language Learning. pp. 107–116. Association for Computational Linguistics (2010)

[6] Davies, J., Hare, J., Samangooei, S., Preston, J., Jain, N., Dupplaw, D., Lewis, P.H.: Identifying the geographic location of an image with a multimodal probability density function. In: MediaEval 2013 / Placing: Geo-coordinate Prediction for Social Multimedia (October 2013), `http://eprints.soton.ac.uk/358836/`

[7] Derczynski, L., Maynard, D., Aswani, N., Bontcheva, K.: Microblog-Genre Noise and Impact on Semantic Annotation Accuracy. In: Proceedings of the 24th ACM Conference on Hypertext and Social Media. ACM (2013)

[8] Gindl, S., Weichselbraun, A., Scharl, A.: Cross-domain contextualisation of sentiment lexicons. In: Proceedings of 19th European Conference on Artificial Intelligence (ECAI-2010). pp. 771–776 (2010)

[9] Go, A., Bhayani, R., , Huang, L.: Twitter sentiment classification using distant supervision. Tech. Rep. CS224N Project Report, Stanford University (2009)

[10] Hare, J., Samangooei, S., Dupplaw, D., Lewis, P.H.: Twitter's visual pulse. In: 3rd ACM International conference on multimedia retrieval. pp. 297–298 (April 2013)

[11] Hare, J.S., Lewis, P.H., Enser, P.G.B., Sandom, C.J.: A linear-algebraic technique with an application in semantic image retrieval. In: Sundaram, H., Naphade, M.R., Smith, J.R., Rui, Y. (eds.) CIVR. Lecture Notes in Computer Science, vol. 4071, pp. 31–40. Springer (2006), `http://dblp.uni-trier.de/db/conf/civr/civr2006.html#HareLES06`

[12] Hare, J.S., Samangooei, S., Dupplaw, D.P.: OpenIMAJ and ImageTerrier: Java libraries and tools for scalable multimedia analysis and indexing of

images. In: Proceedings of the 19th ACM International Conference on Multimedia. pp. 691–694. ACM, New York, NY, USA (2011)
[13] Liebrecht, C., Kunneman, F., van den Bosch, A.: The perfect solution for detecting sarcasm in tweets# not. WASSA 2013 p. 29 (2013)
[14] Liu, X., Zhang, S., Wei, F., Zhou, M.: Recognizing named entities in tweets. In: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies. pp. 359–367 (2011)
[15] Maynard, D., Bontcheva, K., Rout, D.: Challenges in developing opinion mining tools for social media. In: Proceedings of @NLP can u tag #usergeneratedcontent?! Workshop at LREC 2012. Turkey (2012)
[16] Miller, G.A., Beckwith, R., Felbaum, C., Gross, D., Miller, C.Miller, G.A., Beckwith, R., Felbaum, C., Gross, D., Miller, C.Minsky, M.: Five papers on WordNet (1990)
[17] Moghaddam, S., Popowich, F.: Opinion polarity identification through adjectives. CoRR abs/1011.4623 (2010)
[18] Mullaly, A., Gagné, C., Spalding, T., Marchak, K.: Examining ambiguous adjectives in adjective-noun phrases: Evidence for representation as a shared core-meaning. The Mental Lexicon 5(1), 87–114 (2010)
[19] Pak, A., Paroubek, P.: Twitter Based System: Using Twitter for Disambiguating Sentiment Ambiguous Adjectives. In: Proceedings of the 5th International Workshop on Semantic Evaluation. pp. 436–439 (2010), `http://www.aclweb.org/anthology/S10-1097`
[20] Reyes, A., Rosso, P., Veale, T.: A multidimensional approach for detecting irony in twitter. Language Resources and Evaluation pp. 1–30 (2013)
[21] Siersdorfer, S., Hare, J., Minack, E., Deng, F.: Analyzing and predicting sentiment of images on the social web. In: ACM Multimedia 2010. pp. 715–718. ACM (October 2010), `http://eprints.ecs.soton.ac.uk/21670/`
[22] Taboada, M., Brooke, J., Tofiloski, M., Voll, K., Stede, M.: Lexicon-based methods for sentiment analysis. Computational Linguistics 1(September 2010), 1–41 (2011)
[23] Tsur, O., Davidov, D., Rappoport, A.: Icwsm–a great catchy name: Semi-supervised recognition of sarcastic sentences in online product reviews. In: Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media. pp. 162–169 (2010)
[24] Viola, P., Jones, M.: Robust real-time object detection. In: International Journal of Computer Vision (2001)
[25] Wei-ning, W., Ying-lin, Y., Sheng-ming, J.: Image retrieval by emotional semantics: A study of emotional space and feature extraction. In: Systems, Man and Cybernetics, 2006. SMC '06. IEEE International Conference on. vol. 4, pp. 3534–3539 (2006)
[26] Weichselbraun, A., Gindl, S., Scharl, A.: A context-dependent supervised learning approach to sentiment detection in large textual databases. Journal of Information and Data Management 1(3), 329–342 (2010)
[27] Yanulevskaya, V., Van Gemert, J., Roth, K., Herbold, A.K., Sebe, N., Geusebroek, J.M.: Emotional valence categorization using holistic image features. In: Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on. pp. 101–104 (2008)