An Introduction to Concept-Level Sentiment Analysis

Erik Cambria

Temasek Laboratories, National University of Singapore
cambria@nus.edu.sg
http://sentic.net

Abstract. The ways people express their opinions and sentiments have radically changed in the past few years thanks to the advent of social networks, web communities, blogs, wikis, and other online collaborative media. The distillation of knowledge from the huge amount of unstructured information on the Web can be a key factor for marketers who want to create an image or identity in the minds of their customers for their product or brand. These online social data, however, remain hardly accessible to computers, as they are specifically meant for human consumption. The automatic analysis of online opinions, in fact, involves a deep understanding of natural language text by machines, from which we are still very far. To this end, concept-level sentiment analysis aims to go beyond a mere word-level analysis of text and provide novel approaches to opinion mining and sentiment analysis that enable a more efficient passage from (unstructured) textual information to (structured) machine-processable data, in potentially any domain.

Keywords: AI, NLP, concept-level sentiment analysis, big social data analysis.

1 Introduction

Hitherto, online information retrieval has been mainly based on algorithms relying on the textual representation of web pages. Such algorithms are very good at retrieving texts, splitting them into parts, checking the spelling, and counting their words. But when it comes to interpreting sentences and extracting meaningful information, their capabilities are known to be very limited.

Early works aimed to classify entire documents as containing overall positive or negative polarity, or rating scores of reviews. Such systems were mainly based on supervised approaches relying on manually labeled samples, such as movie or product reviews where the opinionist's overall positive or negative attitude was explicitly indicated. However, opinions and sentiments do not occur only at document-level, nor they are limited to a single valence or target. Contrary or complementary attitudes toward the same topic or multiple topics can be present across the span of a document.

Later works adopted a segment-level opinion analysis aiming to distinguish sentimental from non-sentimental sections, e.g., by using graph-based techniques for segmenting sections of a document on the basis of their subjectivity, or by performing a classification based on some fixed syntactic phrases that are likely to be used to express opinions.

F. Castro, A. Gelbukh, M.G. Mendoza (Eds.): MICAI 2013, Part II, LNAI 8266, pp. 478–483, 2013.

In more recent works, text analysis granularity has been taken down to sentence-level, e.g., by using presence of opinion-bearing lexical items (single words or n-grams) to detect subjective sentences, or by exploiting association rule mining for a feature-based analysis of product reviews. These approaches, however, are still far from being able to infer the cognitive and affective information associated with natural language as they mainly rely on knowledge bases that are still too limited to efficiently process text at sentence-level. Moreover, such text analysis granularity might still not be enough as a single sentence may contain different opinions about different facets of the same product or service.

2 Main Approaches to Sentiment Analysis

Existing approaches to sentiment analysis can be grouped into three main categories: keyword spotting, lexical affinity, and statistical methods. Keyword spotting is the most naive approach and probably also the most popular because of its accessibility and economy. Text is classified into affect categories based on the presence of fairly unambiguous affect words like 'happy', 'sad', 'afraid', and 'bored'. The weaknesses of this approach lie in two areas: poor recognition of affect when negation is involved and reliance on surface features. About its first weakness, while the approach can correctly classify the sentence "today was a happy day" as being happy, it is likely to fail on a sentence like "today wasn't a happy day at all". About its second weakness, the approach relies on the presence of obvious affect words that are only surface features of the prose. In practice, a lot of sentences convey affect through underlying meaning rather than affect adjectives. For example, the text "My husband just filed for divorce and he wants to take custody of my children away from me" certainly evokes strong emotions, but uses no affect keywords, and therefore, cannot be classified using a keyword spotting approach.

Lexical affinity is slightly more sophisticated than keyword spotting as, rather than simply detecting obvious affect words, it assigns arbitrary words a probabilistic 'affinity' for a particular emotion. For example, 'accident' might be assigned a 75% probability of being indicating a negative affect, as in 'car accident' or 'hurt by accident'. These probabilities are usually trained from linguistic corpora. Though often outperforming pure keyword spotting, there are two main problems with the approach. First, lexical affinity, operating solely on the word-level, can easily be tricked by sentences like "I avoided an accident" (negation) and "I met my girlfriend by accident" (other word senses). Second, lexical affinity probabilities are often biased toward text of a particular genre, dictated by the source of the linguistic corpora. This makes it difficult to develop a reusable, domain-independent model.

Statistical methods, such as Bayesian inference and support vector machines, have been popular for affect classification of texts. By feeding a machine learning algorithm a large training corpus of affectively annotated texts, it is possible for the system to not only learn the affective valence of affect keywords (as in the keyword spotting approach), but also to take into account the valence of other arbitrary keywords (like lexical affinity), punctuation, and word co-occurrence frequencies. However, traditional statistical methods are generally semantically weak, meaning that, with the exception of obvious affect keywords, other lexical or co-occurrence elements in a statistical model have little predictive value individually. As a result, statistical text classifiers only work with acceptable accuracy when given a sufficiently large text input. So, while these methods may be able to affectively classify user's text on the page- or paragraph- level, they do not work well on smaller text units such as sentences or clauses.

3 Concept-Level Sentiment Analysis

Concept-based approaches to sentiment analysis focus on a semantic analysis of text through the use of web ontologies or semantic networks, which allow the aggregation of conceptual and affective information associated with natural language opinions. By relying on large semantic knowledge bases, such approaches step away from blind use of keywords and word co-occurrence count, but rather rely on the implicit features associated with natural language concepts. Unlike purely syntactical techniques, concept-based approaches are able to detect also sentiments that are expressed in a subtle manner, e.g., through the analysis of concepts that do not explicitly convey any emotion, but which are implicitly linked to other concepts that do so.

The analysis at concept-level is intended to infer the semantic and affective information associated with natural language opinions and, hence, to enable a comparative fine-grained feature-based sentiment analysis. Rather than gathering isolated opinions about a whole item (e.g., iPhone5), users are generally more interested in comparing different products according to their specific features (e.g., iPhone5's vs Galaxy S3's touchscreen), or even sub-features (e.g., fragility of iPhone5's vs Galaxy S3's touchscreen). In this context, the construction of comprehensive common and common-sense knowledge bases is key for feature-spotting and polarity detection, respectively [1]. Common-sense, in particular, is necessary to properly deconstruct natural language text into sentiments— for example, to appraise the concept "small room" as negative for a hotel review and "small queue" as positive for a post office, or the concept "go read the book" as positive for a book review but negative for a movie review [2].

Current approaches to concept-level sentiment analysis mainly leverage on existing affective knowledge bases such as ANEW [3], WordNet-Affect [4], ISEAR [5], SentiWordNet [6], and SenticNet [7]. In [8], for example, a concept-level sentiment dictionary is built through a two-step method combining iterative regression and random walk with in-link normalization. ANEW and SenticNet are exploited for propagating sentiment values based on the assumption that semantically related concepts share common sentiment. Moreover, polarity accuracy, Kendall distance, and average-maximum ratio are used, in stead of mean error, to better evaluate sentiment dictionaries. A similar approach is adopted in [9], which presents a methodology for enriching SenticNet concepts with affective information by assigning an emotion label to them. Authors use various features extracted from ISEAR, as well as similarity measures that rely on the polarity data provided in SenticNet (those based on WordNet-Affect) and ISEAR distance-based measures, including point-wise mutual information, and emotional affinity. Another recent work that builds upon an existing affective knowledge base is [10], which proposes the re-evaluation of objective words in SentiWordNet by assessing the sentimental relevance of such words and their associated sentiment sentences. Two sampling strategies are proposed and integrated with support vector machines for sentiment classification. According to the experiments, the proposed approach significantly outperforms the traditional sentiment mining approach, which ignores the importance of objective words in SentiWordNet. In [11], the main issues related to the development of a corpus for opinion mining and sentiment analysis are discussed both by surveying the existing works in this area and presenting, as a case study, an ongoing project for Italian, called Senti–TUT, where a corpus for the investigation of irony about politics in social media is developed.

Other works explore the ensemble application of knowledge bases and statistical methods. In [12], for example, a hybrid approach to combine lexical analysis and machine learning is proposed in order to cope with ambiguity and integrate the context of sentiment terms. The context-aware method identifies ambiguous terms that vary in polarity depending on the context and stores them in contextualized sentiment lexicons. In conjunction with semantic knowledge bases, these lexicons help ground ambiguous sentiment terms to concepts that correspond to their polarity. More machine-learning based works include [13], which introduces a new methodology for the retrieval of product features and opinions from a collection of free-text customer reviews about a product or service. Such a methodology relies on a language-modeling framework that can be applied to reviews in any domain and language provided with a seed set of opinion words. The methodology combines both a kernel-based model of opinion words (learned from the seed set of opinion words) and a statistical mapping between words to approximate a model of product features from which the retrieval is carried out.

Other recent works in the context of concept-level sentiment analysis include tasks such as domain adaptation [14], opinion summarization [15], and multimodal sentiment analysis [16,17]. In the problem of domain adaptation, there are two distinct needs, namely labeling adaptation and instance adaptation. However, most of current research focuses on the former attribute, while neglects the latter one. In [14], a comprehensive approach, named feature ensemble plus sample selection (SS-FE), is proposed. SS-FE takes both types of adaptation into account: a feature ensemble (FE) model is first adopted to learn a new labeling function in a feature re-weighting manner, and a PCA-based sample selection (PCA-SS) method is then used as an aid to FE. A first step towards concept-level summarization is done by STARLET [15], a novel approach to extractive multi-document summarization for evaluative text that considers aspect rating distributions and language modeling as summarization features. Such features encourage the inclusion of sentences in the summary that preserve the overall opinion distribution expressed across the original reviews and whose language best reflects the language of reviews. The proposed method offers improvements over traditional summarization techniques and other approaches to multi-document summarization of evaluative text.

A sub-field of sentiment analysis that is becoming increasingly popular is multimodal sentiment analysis. [16], for example, considers multimodal sentiment analysis based on linguistic, audio, and visual features. A database of 105 Spanish videos of 2 to 8 minutes length containing 21 male and 84 female speakers was collected randomly from the social media website YouTube and annotated by two labellers for ternary sentiment. This led to 550 utterances and approximately 10,000 words. The authors state that the data is available per request. The joint use of the three feature types leads to a significant improvement over the use of each single modality. This is further confirmed on another set of English videos. In [17], instead, authors introduce the ICT-MMMO database of personal movie reviews collected from YouTube (308 clips) and ExpoTV (78 clips). The final set contains 370 of these 1-3 minutes English clips in ternary sentiment annotation by one to two coders. The feature basis is formed by 2 k audio features, 20 video features, and different textual features for selection. Then, different levels of domain-dependence are considered: in-domain analysis, cross-domain analysis based on the 100 k textual Metacritic movie review corpus for training, and use of on-line knowledge sources. This shows that cross-corpus training works sufficiently well, and language-independent audiovisual analysis to be competitive with linguistic analysis.

4 Conclusion

Between the dawn of civilization through 2003, there were just a few dozens exabytes of information on the Web. Today, that much information is created weekly. The advent of the Social Web has provided people with new tools for creating and sharing, in a time and cost efficient way, their own contents, ideas, and opinions with virtually the millions of people connected to the World Wide Web. The opportunity to capture the opinions of the general public about social events, political movements, company strategies, marketing campaigns, and product preferences has raised increasing interest both in the scientific community, for the exciting open challenges, and in the business world, for the remarkable fallouts in marketing and financial market prediction. This huge amount of useful information, however, is mainly unstructured as specifically produced for human consumption and, hence, it is not directly machine-processable.

Concept-level sentiment analysis can help with this in which, unlike other word-based approaches, it focuses on a semantic analysis of text through the use of web ontologies or semantic networks, which allow the aggregation of conceptual and affective information associated with natural language opinions. The validity of concept-based approaches, however, depends on the depth and breadth of the employed knowledge bases. Without a comprehensive resource that encompasses human knowledge, it is not easy for an opinion mining system to grasp the semantics associated with natural language text. Another limitation of semantic approaches is in the typicality of their knowledge bases. Knowledge representation, in fact, is usually strictly defined and does not allow different concept nuances to be handled, as the inference of semantic and affective features associated with concepts is bounded by the fixed, flat representation. In the big social data context, finally, semantic parsing techniques will be key in quickly identifying natural language concepts from free text without requiring time-consuming phrase structure analysis [18].

References

1. Cambria, E., Song, Y., Wang, H., Howard, N.: Semantic multi-dimensional scaling for open-domain sentiment analysis. IEEE Intelligent Systems (2013), doi:10.1109/MIS.2012.118
2. Cambria, E., Schuller, B., Xia, Y., Havasi, C.: New avenues in opinion mining and sentiment analysis. IEEE Intelligent Systems 28(2), 15–21 (2013)

3. Bradley, M., Lang, P.: Affective norms for english words (ANEW): Stimuli, instruction manual and affective ratings. Technical report, The Center for Research in Psychophysiology, University of Florida (1999)
4. Strapparava, C., Valitutti, A.: WordNet-Affect: An affective extension of WordNet. In: LREC, Lisbon, pp. 1083–1086 (2004)
5. Bazzanella, C.: Emotions, language and context. In: Weigand, E. (ed.) Emotion in Dialogic Interaction. Advances in the complex, pp. 59–76, Benjamins, Amsterdam (2004)
6. Esuli, A., Sebastiani, F.: SentiWordNet: A publicly available lexical resource for opinion mining. In: LREC (2006)
7. Cambria, E., Havasi, C., Hussain, A.: SenticNet 2: A semantic and affective resource for opinion mining and sentiment analysis. In: FLAIRS, Marco Island, pp. 202–207 (2012)
8. Tsai, A., Tsai, R., Hsu, J.: Building a concept-level sentiment dictionary based on commonsense knowledge. IEEE Intelligent Systems 28(2), 22–30 (2013)
9. Poria, S., Gelbukh, A., Hussain, A., Das, D., Bandyopadhyay, S.: Enhanced SenticNet with affective labels for concept-based opinion mining. IEEE Intelligent Systems 28(2), 31–30 (2013)
10. Hung, C., Lin, H.K.: Using objective words in SentiWordNet to improve sentiment classification for word of mouth. IEEE Intelligent Systems 28(2), 47–54 (2013)
11. Bosco, C., Patti, V., Bolioli, A.: Developing corpora for sentiment analysis and opinion mining: A survey and the Senti-TUT case study. IEEE Intelligent Systems 28(2), 55–63 (2013)
12. Weichselbraun, A., Gindl, S., Scharl, A.: Extracting and grounding context-aware sentiment lexicons. IEEE Intelligent Systems 28(2), 39–46 (2013)
13. García-Moya, L., Anaya-Sanchez, H., Berlanga-Llavori, R.: A language model approach for retrieving product features and opinions from customer reviews. IEEE Intelligent Systems 28(3), 19–27 (2013)
14. Xia, R., Zong, C., Hu, X., Cambria, E.: Feature ensemble plus sample selection: A comprehensive approach to domain adaptation for sentiment classification. IEEE Intelligent Systems 28(3), 10–18 (2013)
15. Di Fabbrizio, G., Aker, A., Gaizauskas, R.: Summarizing on-line product and service reviews using aspect rating distributions and language modeling. IEEE Intelligent Systems 28(3), 28–37 (2013)
16. Perez-Rosas, V., Mihalcea, R., Morency, L.P.: Multimodal sentiment analysis of Spanish online videos. IEEE Intelligent Systems 28(3), 38–45 (2013)
17. Wollmer, M., Weninger, F., Knaup, T., Schuller, B., Congkai, S., Sagae, K., Morency, L.P.: YouTube movie reviews: In, cross, and open-domain sentiment analysis in an audiovisual context. IEEE Intelligent Systems 28(3), 46–53 (2013)
18. Cambria, E., Rajagopal, D., Olsher, D., Das, D.: Big social data analysis. In: Akerkar, R. (ed.) Big Data Computing, pp. 401–414. Chapman and Hall/CRC (2013)