Design Engineering
ISSN: 0011-9342 | Year 2021
Issue: 8 | Pages: 14959 - 14964

Current Challenges and New Directions in Sentiment Analysis on Web based media Information

K.Brindha
Research Scholar
NGM College
Pollachibrindharajmohan004@gmail.com

Dr.E.Ramadevi
Associate Professor
NGM College,Pollachi
ramasai1970@gmail.com

Abstract

Numerous organizations are utilizing online media organizations to convey various administrations and interface with customers and gather data about the contemplations and perspectives on people. Sentiment Analysis or Opinion Mining is a method of AI that detects polarities like positive or negative contemplations inside the text, full records, passages, lines, or subsections.

In this part, we point out a portion of the issues and challenges that actually remain, questions that have not been investigated adequately, and new issues rising up out of taking on new feeling examination issues. The purpose of this paper is to explain the numerous tasks and challenges of Opinion Mining and Sentiment Analysis.

Keywords—Sentiment, Polarity, Opinion mining

I. INTRODUCTION

Sentiment analysis assesses the sentiment polarity orientation of each opinionated word or phrase, whether it is positive, negative, or neutral. It expresses a writer's or speaker's opinion in a concise manner. Sentiment analysis can be performed on a word, phrase, or document level.A portion of the more aggressive issues that need more work incorporate distinguishing sentiments at different degrees of message granularities (terms, sentences, passages, and so forth); identifying slant of the peruser or conclusion of substances referenced in the message; recognizing notion towards parts of items; recognizing position towards pre-determined focuses on that may not be unequivocally referenced in the message and that may not be the objectives of assessment in the message; and identifying semantic jobs of estimation. Since numerous notion examination frameworks depend on feeling vocabularies, we talk about capacities and restrictions of existing physically and consequently made conclusion dictionaries. we examine the troublesome issue of notion creation—how to foresee the opinion of a blend of terms. All the more explicitly, we examine the assurance of estimation of expressions (that might incorporate negators, degree verb modifiers, and intensifiers) and sentiment of sentences and tweets

We discuss challenges in annotation of data for sentiment. We provide categories of sentences that are particularly challenging for sentiment annotation. This is followed by a discussion on the challenges of applying sentiment analysis to downstream applications, and finally, some concluding remarks.

II. SENTIMENT ANALYSIS TASKS

A. Sentiment at Different Text Granularities

Sentiment can be determined at various levels: from sentiment associations of words and phrases; to sentiment of sentences, SMS messages, chat messages, and tweets; to sentiment in product reviews, blog posts, and whole documents

Amazing -positive

adulterate -negative

Paper - neutral

These lexicons can be made either by manual explanation or through programmed implies. Physically made lexicons will in general be in the request for a couple thousand passages, however automatically produced vocabularies can catch notion relationship for many thousands unigrams (single word strings) and in any event, for bigger articulations like bigrams (two-word arrangements) and trigrams (three-word groupings). Passages in a naturally produced lexsymbol frequently likewise incorporate a genuine esteemed score showing the strength of relationship between the word and the valence class. While slant dictionaries are regularly helpful in sentence-level slant analysis1, similar terms might pass on various notions in various settings.

These lexicons can be created either by manual annotation or through automatic means. Manually created lexicons tend to be in the order of a few thousand entries, but automatically generated lexicons can capture sentiment associations for hundreds of thousands unigrams (single word strings) and even for larger expressions such as bigrams (two-word sequences) and trigrams(three-word sequences). Entries in an automatically generated lexicon often also include a real-valued score indicating the strength ofassociation between the word and the valence category. While sentiment lexicons are often useful in sentence-level sentiment analysis1, the same terms may convey different sentiments in different contexts.

B. Detecting Affect and Emotions

Sentiment analysis is most commonly used to refer to the goal of determining the valence or polarity of a piece of text. However, it can refer more generally to determining one’s attitude towards a particular target or topic. Here, attitude can even mean emotional or affectual attitude such as frustration, joy, anger, sadness, excitement, and so on. Russell (1980) developed a circumplex model of affect and showed that it can be characterized by two primary dimensions: valence (positive and negative dimension) and arousal (degree of reactivity to stimulus). Thus, it is not surprising that large amounts of work in sentiment analysis is focused on determining valence. However, there is barely any work on automatically detecting arousal and a relatively small amount of work on detecting emotions such as anger, frustration, sadness, and optimism (Strapparava & Mihalcea, 2007; Aman & Szpakowicz, 2007; Tokuhisa, Inui, & Matsumoto, 2008; Neviarouskaya, Prendinger, & Ishizuka, 2009; Bellegarda, 2010)

C.*Sentiment of Phrases, Sentences, and Tweets: Sentiment Composition*

Semantic composition, which aims at determining a representation of the meaning of two words through manipulations of their individual representations, has gained substantial attention in recent years with work from Mitchell and Loapata (2010), Baroni and Zamparelli (2010), Rudolph and Giesbrecht (2010), Yessenalina and Cardie (2011), Grefenstette, Dinu, Zhang, Sadrzadeh, and Baroni (2013), Grefenstette and Sadrzadeh (2011), and Turney (2014). Socher, Huval, Manning, and Ng (2012) and Mikolov, Sutskever, Chen, Corrado, and Dean (2013) introduced deep learning models and distributed word representations in vector space

ISSN: 0011-9342 | Year 2021
Issue: 8 | Pages: 14959 - 14964

(word embeddings) to obtain substantial improvements over the state-of-the- art in semantic composition.

D.*Negated Expressions*

Morante and Sporleder (2012) define negation to be a grammatical category that allows the changing of the truth value of a proposition". Negation is often expressed through the use of negative signals or negator words such as not and never, and it can significantly affect the sentiment of its scope. Understanding the impact of negation on sentiment improves automatic analysis of sentiment. Earlier works on negation handling employ simple heuristics such as flipping the polarity of the words in a negator's scope (Kennedy & Inkpen, 2005; Choi & Cardie, 2008) or changing the degree of sentiment of the modified word by a fixed constant (Taboada, Brooke, Tofiloski, Voll, & Stede, 2011). Kiritchenko et al. (2014b) capture the impact of negation by creating separate sentiment lexicons for words seen in affirmative context and those seen in negated contexts. They use a hand-chosen list of negators and determine scope to be starting from the negator and ending at the first punctuation (or end of sentence).

Several aspects about negation are still not understood though: for example, in what contexts does the same negator impact the sentiment of its scope more and in what contexts is the impact less; how do people in different communities and cultures use negations differently; and how negations of sentiment expressions should be dealt with by paraphrase and textual entailment systems.

E.Sentiment of Sentences, Tweets, and SMS messages

Socher et al. (2013) proposed a word-embeddings based model that learns the sentiment of term compositions. They obtain state-of-the-art results in determining both the overall sentiment and sentiment of constituent phrases in movie review sentences. This has inspired tremendous interest in more embeddings-based work for sentiment composition (Dong, Wei, Zhou, & Xu, 2014; Kalchbrenner,Grefenstette, & Blunsom, 2014). These recursive models do not require any hand- crafted features or semantic knowledge, such as a list of negation words or sentiment lexicons. However, they are computationally intensive and need substantial additional annotations (word and phrase- level sentiment labeling). Nonetheless,use of word-embeddings in sentiment composition is still in its infancy, and we will likely see much more work using these techniques in the future.

F.Sentiment in Figurative Expressions

There is growing interest in detectingfigurative language, especially irony and sarcasm(Carvalho, Sarmento, Silva, & De Oliveira, 2009; Reyes, Rosso, & Veale, 2013; Veale & Hao, 2010; Filatova, 2012; Gonz´alez-Ib´anez, Muresan, & Wacholder, 2011). In 2015, a SemEval shared task was organized on detecting sentiment in tweets rich in metaphor and irony 6 Participants were asked to determine the degree of sentiment for each tweet where the score is a real number in the range from -5 (most negative) to +5 (most positive). One of the characteristics of the data is that a large majority is negative; thereby suggesting that ironic tweets are largely negative.

III. CHALLENGESINAPPLYINGSENTIMENTANALYSIS

Sentiment analysis is commonly applied in several areas including tracking sentiment towards products, movies, politicians, and companies (O'Connor, Balasubramanyan, Rout- ledge, & Smith, 2010; Pang & Lee, 2008), improving customer relation models (Bougie, Pieters, & Zeelenberg, 2003), detecting happiness and well-being (Schwartz, Eichstaedt,

ISSN: 0011-9342 | Year 2021
Issue: 8 | Pages: 14959 - 14964

Kern, Dziurzynski, Lucas, Agrawal, Park, et al., 2013), tracking the stock market (Bollen, Mao, & Zeng, 2011), and improving automatic dialogue systems (Vel´asquez, 1997; Ravaja, Saari, Turpeinen, Laarni, Salminen, & Kivikangas, 2006).

The use of sentiment analysis to harvest massive amounts of data has become a major academic topic. This document outlines some of the most often utilised sentiment analysis challenges and obstacles. Now, commercial groups and academia are collaborating to develop the finest sentiment analysis technology. Although various algorithms have been utilised in sentiment analysis and have produced positive results, no algorithm can tackle all of the problems. The classification of sentiment is found to be domain dependent.

A. Sentiment Analysis in Twitter

(Rosenthal et al., 2014a) had a separate test set involving sarcastic tweets. Participants were asked not to train their system on sarcastic tweets, but rather apply their regular sentiment system on this new test set; the goal was to determine performance of regular sentiment systems on sarcastic tweets. It was observed that the performances dropped by about 25 to 70 percent, thereby showing that systems must be adjusted if they are to be applied to sarcastic tweets.

- Some sentence types that are especially challenging for sentiment annotation (using either the simple sentiment questionnaire or the semantic-role based sentiment questionnaire) are listed below:
- *Speaker's emotional state*: The speaker's emotional state may or may not have the same polarity as the opinion expressed by the speaker. For example, a politician's tweet can imply both a negative opinion about a rival's past indiscretion, and a joyous mental state as the news will impact the rival adversely.
- *Success or failure of one side w.r.t. another:*

Often sentences describe the success or failure of one side w.r.t. another side—for example-Australia won Finland in the match - This is not because the *speakers* were expressing negative opinion towards the Finland team, but rather simply because Australia was the focus of attention and traditionally Australian teams have been strong.

- *Neutralreportingofvalencedinformation:*

If the speaker does not give any indication of her own emotional state but describes valenced eventsor situations, then it is unclear whether to consider these statements as neutral unemotional reporting of developments or whether to assume that the speaker is in a negative emotional state (sad, angry, etc.). Example:The war has created millions of refugees.

- *Sarcasm and ridicule:* From the standpoint of assigning a single label of feeling, sarcasm and ridicule are challenging since they can often imply a good emotional state of the speaker (joy from mocking someone or something) even if they have a negative attitude about that person or thing.
- *Quoting somebody else or re-tweeting*: Quotes and retweets are challenging to annotate for sentiment because it is frequently ambiguous and not explicitly stated whether the person who quotes (or retweets) shares the same viewpoints as the quotee.
- *Spam and false reviews detection:* The internet contains both genuine and spam content. This spam content should be removed before processing for successful Sentiment categorization. This can be accomplished by looking for duplicates, spotting outliers, and assessing the reviewer's reputation. Some methods for detecting bogus comments have surfaced. Semantic similarity and grammatical analysis, for example. These methods can detect bogus comments, although their accuracy is low. Bogus comments can have either positive or negative sentiments

• *Challenges in Multilingual Sentiment Analysis*

Some of the less-explored areas in the field of multilingual sentiment analysis include: how to translate text so that the degree of sentiment in the source text is preserved; how to translate text so that the degree of sentiment in the source text is preserved; and how to translate text so that the degree of sentiment in the source text is preserved. how sentiment modifiers such as negators and modals function differently in different languages; how automatic and human translations differ in terms of sentiment; and how to translate figurative language

• *Domain independence:* The domain dependent character of sentiment words is the major hurdle for opinion mining and sentiment analysis. In one area, one collection of features may perform exceptionally well, but in another domain, it may perform horribly.

IV. Conclusion

Different types of classification algorithms should be coupled to overcome theunique flaws, profit from one another's strengths, and improve sentiment classification performance. Such apps are in high demand in the market since every company wants to know how customers feel about their products and services, as well as those of their competitors. New applications for sentiment analysis can be created. A future challenge in applying sentiment classification approaches and tools for sentiment analysis of social media posts is overcoming the ambiguity that actually represents a specific problem because coreference information is not easily accessible. Irony and sarcasm are common in the posts examined, but they can be difficult to detect.To overcome this limitation, an evolution of approaches and tools is required.Although sentiment analysis techniques and algorithms have progressed, many difficulties in this subject remain unsolved. More research can be done in the future to address these issues.

References:

[1] B. Liu, Sentiment Analysis and Opinion Mining, Morgan & Claypool, San Rafael, CA, 2012.
[2] NLP market https://www.tractica.com/newsroom/press-releases/natural-language-processing-market-to-reach-22–3-billion-by-2025/.
[3] Kim, Y. Convolutional neural networks for sentence classification. arXiv 2014, arXiv:1408.5882.
[4] Nal, K.; Grefenstette, E.; Blunsom, P. A convolutional neural network for modeling sentences. arXiv 2014, arXiv:1404.2188.
[5] Lei, T.; Barzilay, R.; Jaakkola, T. Molding cnns for text: Non-linear, non-consecutive convolutions. arXiv 2015, arXiv:1508.04112.
[6] E. Boiy, P. Hens, K. Deschacht, and M.-F. Moens, "Automatic sentiment analysis in on-line text," in Proceedings of the 11th International Conference on Electronic Publishing, pp. 349–360, 2007.
[7] J. Kamps, M. Marx, R. J. Mokken, and M. De Rijke, "Using wordnet to measure semantic orientations of adjectives," 2004.
[8] C. Fellbaum, "Wordnet: An electronic lexical database (language, speech, and communication)," 1998.
[9] D. Pucci, M. Baroni, F. Cutugno, and A. Lenci, "Unsupervised lexical substitution with a word space model," in Proceedings of EVALITA workshop, 11th Congress of Italian Association for Artificial Intelligence, Citeseer, 2009.
[10] A. Balahur, J. M. Hermida, and A. Montoyo, "Building and exploiting emotinet, a knowledge base for emotion detection based on the appraisal theory model," Affective Computing, IEEE Transactions on, vol. 3, no. 1, pp. 88–101, 2012.

[11] C. Strapparava and R. Mihalcea, “Semeval-2007 task 14: Affective text,” in Proceedings of the 4th International Workshop on Semantic Evaluations, pp. 70–74, Association for Computational Linguistics, 2007.
[12] G. Vinodhini and R. Chandrasekaran, “Sentiment analysis and opinion mining: A survey,” International Journal, vol. 2, no. 6, 2012.
[13] Y. Mejova, “Sentiment analysis: An overview,” Comprehensive exam paper, available on http://www. cs. uiowa. edu/˜ ymejova/publications/CompsYelenaMejova. pdf [2010-02-03], 2009.
[14] A. Pak and P. Paroubek, “Twitter as a corpus for sentiment analysis and opinion mining,” in Proceedings of LREC, vol. 2010, 2010.
[15] R. Xia, C. Zong, and S. Li, “Ensemble of feature sets and classification algorithms for sentiment classification,” Information Sciences: an International Journal, vol. 181, no. 6, pp. 1138–1152, 2011.