RUDIMENTARY LEXICON BASED METHOD FOR SARCASM DETECTION

Peter Clews
University of Worcester
UK
p.clews@worc.ac.uk

Dr. Joanne Kuzma
University of Worcester
UK
j.kuzma@worc.ac.uk

ABSTRACT

The purpose of this paper is to establish if rudimentary methods can be used for classifying text as being sarcastic using data taken from the social media website, Twitter. Data collection for this study was carried out using text extracted from Twitter. It applies string matching against positive sentiment and interjection lexicons to test if the presence of both can be used to classify content as being sarcastic. The result shows that the most frequently used terms are too generic to be suitable for a sarcasm specific lexicon. It further shows that Boolean matches to two lexicons are suitable for classification of text as being sarcastic. This is significant as many methods require significant time in collecting and analysing the data to be used within the classification process, as well as complex algorithms to conduct the task. By using simplistic processes, it is hoped that some of the challenges related to performance are overcome. Additionally, this study is the precursor to planned further research into sarcasm detection methods.

Keywords: Sentiment analysis, Twitter, sarcasm detection, positive sentiments, tweets.

INTRODUCTION

This research seeks to determine if rudimentary methods of text mining and sentiment analysis can be used to classify text as being sarcastic. This is significant as many methods require significant time in collecting and analysing the data to be used within the classification process, as well as complex algorithms to conduct the task. By using simplistic processes, it is hoped that some of the challenges related to performance are overcome. The focus is to explore the use of positive sentiments and interjections to be able to classify short text as containing sarcasm. It combines features from three different studies into sarcasm classification, which were then used for the creation of two lexicons.

Riloff et al (2014) state that one of the most common forms of sarcasm contain positive sentiment and a negative task. However, whilst it was agreed that most sarcastic utterances contain positive sentiment, it was not agreed that all sarcasm relates to a negative activity. As a result, the negative task aspect was replaced for this study. Kovaz et al (2013) and Kreuz and Caucci (2007) found that most sarcastic text contained interjections, such as “wow!” It was agreed that this is a reasonable feature within most sarcastic texts and so was used as a second feature within this study.

For this study, if sarcasm is to be declared then the presence of both positive sentiment and interjections within the text is needed. There are two research questions which this study seeks to answer:

1. Is it possible to identify sarcasm using positive sentiment and interjections?
2. Can a string matching method against two lexicons produce accurate results?

LITERATURE REVIEW

Challenges

Researchers and organisations were quick to discover value within all the information and tried to organise the content. Originally, much of this looked into trying to classify the data by topic, such as news or sports. However, it became apparent that much of the data contained the opinions of the user towards a product, service or event and a topical classification overlooked a, potentially, valuable resource. Pang et al (2002) and Turney (2002) were among the first to seek a method of classifying the text based on those opinions of the users. This has become to be known as either sentiment analysis or opinion mining and is considered to be a natural language processing activity. The aim is to provide a classification being based on the polarity of opinions towards a given domain; either positive, negative or neutral.

Sentiment analysis tasks can be conducted using a lexicon dictionary of scored words which are considered to be positive or negative, using machine learning algorithms or by combining the two methods. Lexicons can be obtained by using existing, well established and freely available examples, such as SentiWordNet, or created. They can be created either manually or by using seed words. The machine learning algorithms require a set of data which has been labelled with the possible outcomes (Fernández -Gavilanes et al, 2016).

One of the greatest challenges with conducting a sentiment analysis task is how to accurately identify, and adjust for, text which is either sarcastic or ironic. As a form of verbal irony, sarcasm is used to convey the opposite sentiment of what is literally stated (Filik et al, 2016). The detection of sarcasm or irony within written passages is troublesome in comparison to detecting it in speech (Thompson & Filik, 2016). This is because when sarcasm is used in spoken communications there may be a number of additional non-verbal indicators, such as tone of voice or gesticulations. However, there are often none of these additional markers present in written communications.

Perspectives of sarcasm

Several studies have reviewed the lexical basis for sarcasm. Kreuz and Caucci (2007) successfully sought to establish whether there were lexical influences which would impact on the perception of sarcasm, even without the full context. They found the use of interjections within the text was a substantial predictor of classification outcomes. González-Ibánez et al (2011) reviewed the lexical and pragmatic factors for separating sarcastic tweets from those expressing positive or negative sentiments. They found that the lexical factors alone were insufficient at identifying the distinction between sarcastic and sentiment containing tweets. Kovaz et al (2013) supports the findings of Kreuz and Caucci (2007), that there are certain aspects, such as interjections that can be utilised to indicate the presence of sarcasm.

Regardless of which theoretical framework is followed, the research shows that both context and lexical features are important for the detection of sarcasm within a corpus. Both Tsur et al (2010) and González-Ibánez et al (2011) found that context is an important consideration for determining sarcasm. Given the example statement "nice hair", the sentiment could be markedly different depending on the context in which it applies. If it is meant as a complement to someone who has a particularly pleasing hair style, then the message would be very supportive. However, if the utterance is aimed at someone with an unflattering hair style then the message would clearly be sarcastic (Tsur et al, 2010).

Approaches to sarcasm detection

There are generally two types of approaches for the automated detection of sarcasm. These are rule-based and statistical approaches (Joshi et al, 2016). These, usually, seek to utilise novel features or certain patterns to classify text as being sarcastic. With the popularity of Twitter, there arose hashtag-based approaches; however, the hashtag-based approach is classed as being one of the rule-based methods.

Maynard and Greenwood (2014) utilised a hashtag-based approach when looking at a corpus of Tweets. Using the basis that certain hashtags can be used to denote that the message is sarcastic in nature, Maynard and Greenwood (2014) sought to examine the impact sarcasm has on sentiment analysis. They found that the detection of sarcasm in tweets has little impact on the accuracy of a sentiment analysis task. Additionally, they state that rules needed to be applied to the hashtags in order to increase the accuracy. Bharti et al (2015) produced two rule-based algorithms to detect sarcasm. The first is a parsed-based lexicon generation algorithm which detects situation phases that contain sentiment. Using this algorithm, sarcasm is classified when a negative word or phrase is detected within a positive message. Veale and Hao (2010) use a wildcard search, in the form of "as * as a *", on the Google search engine to determine if a given simile is intended to be sarcastic. They proposed a nine-step process for classifying the results as being either sarcastic or non-sarcastic.

Tsur et al's (2010) Semi-supervised Algorithm for Sarcasm Identification (SASI) is considered a statistical method (Joshi et al, 2016). This is due to the use of pattern recognition to identify features of a corpus which has been pre-labelled as being sarcastic. These features are converted to a numeric value to indicate if there was an exact, partial or no match. González-Ibánez et al (2011) used a 'bag-of-words' approach with the addition of emoticons as a feature. They compared the performance of both machine learning algorithms with human judges and found that neither performed particularly well in identifying sarcasm.

Sources of text data

Sarcasm is not just limited to social media or other forms of user generated content on the internet. There are many different sources for text data which contain sarcastic messages or verses. These data sources can be classed as being short text (such as tweets), long text (product reviews), or more novel sources (such as telephone transcripts) (Joshi et al, 2016).
Social media permits access to several different forms of data (Joshi et al, 2012). Owing to the popularity of Twitter, the ease of use of their API and the presence of hashtags, many studies have looked to use this data as a corpus (Riloff et al, 2013; Maynard & Greenwood, 2014; Ptácek et al, 2014; Davidov et al, 2010). Owing to the fact that the author is the only person who can label the text with a hashtag, several studies use this feature. Davidov et al (2010), González-Ibánez et al (2011) and Reyes et al (2013) all used the hashtag as a means of identifying their datasets, with all of them using the #sarcasm hashtag. Whilst Twitter is a popular source of data, another that has been utilised is Reddit (Joshi et al, 2016).

Longer texts taken from discussion boards and product reviews are also featured in the studies into sarcasm detection. Lukin and Walker (2013) used a dataset taken from the Internet Argument Corpus, a freely available corpus of discussion board posts. This dataset has several different labels, Lukin and Walker (2013) focused on those labelled as being sarcastic in the hopes of testing a bootstrapping method to improve accuracy of precision and recall. Reyes and Rosso (2014) used a dataset consisting of a mixture of movie reviews, book reviews and news articles which had been with both sarcasm and sentiment. Reyes and Rosso (2012) used viral product reviews taken from Amazon in order to identify key characteristics which represent irony. Filatova (2012) also used product reviews found on Amazon with a view of identifying both document and sentence levels of sarcasm. Another study to use product reviews taken from Amazon was Buschmeier et al (2014). The used a total of one-thousand, two-hundred and fifty-four reviews and found that four-hundred and thirty-seven to be sarcastic. One of the largest corpus' to be taken from Amazon was by Tsur et al (2010) who used sixty-six-thousand reviews in the creation of the SASI algorithm.

A number of novel data sources have also been used for the detection of the presence of sarcasm in text (Joshi et al, 2016). Tepperman et al (2006) used a dataset consisting of one-hundred and thirty-one call centre transcripts. They used the presence of 'yeah right' as a marker for the presence of sarcasm. Kreuz and Caucci (2007) used twenty excerpts from novels, where the author has stated sarcasm was present, to identify the lexical markers of sarcasm. Veale and Hao (2010) used Google to create a dataset which uses the text pattern 'as * as a *'. This process resulted in twenty-thousand similes which could be markers for sarcasm or not.

Issues with sarcasm detection

Issues surrounding the annotation of sarcastic text can be divided into several sections. The first relates to how hashtags are used in supervised methods. Many of the studies mentioned above used the hashtag features as a means of identifying sarcastic intent. However, Liebrecht et al (2013) found that whilst the hashtags are used to select data to use they are frequently removed during pre-processing. To Liebrecht et al (2013), the use of the hashtags are the extra linguistic versions of non-verbal expressions present in verbal exchanges of sarcasm. Joshi et al (2016) state that the hashtag itself is, frequently, the only marker for sarcastic intent within a message and therefore its removal during pre-processing alters the messages meaning. The second area identified by Joshi et al (2016) concerns the data imbalance. This is due to sarcasm being a rare phenomenon and can result in only a small fraction of the data set being sarcastic. An example of this is Barbieri et al (2014) who, from a corpus of twenty-five thousand, four-hundred and fifty records, found only twelve percent to be sarcastic. Lui et al (2014) created a 'novel multi-strategy ensemble learning approach' to account for this issue. Testing this on six datasets, three English and three Chinese, they found that by accommodating for the imbalances within the dataset produced greater degrees of accuracy in sarcasm detection.

METHODOLOGY

All the data for this study was taken from Twitter using their Application Programming Interface (API. Twitter was chosen as the source due to its characteristics. Mainly owing to it having a limit on the amount of characters a user can write which allows for very short messages for them to convey their opinions. Another important feature in the decision to use Twitter, as a corpus for this study, was the inclusion of hashtags as a means of users self-classifying content. This was crucial as the API permits data to be selected via a search on the hashtags.

Using the search feature, two corpora were created, both of which were restricted to English language only. One, containing five-thousand records which contained the hashtag #sarcasm and a second containing one-thousand records which used the hashtag #trump. The use of the

ISSN 2309-0405

#sarcasm hashtag was both important and convenient as it allows the authors of the tweets to mark their message as being sarcastic. The #trump hashtag relates to Donald Trump, the businessman and Republican presidential candidate (at time of research). This was chosen due to the widespread media coverage during recent months and the polarities of opinions people have towards Donald Trump's campaigning to become president of the United States of America.

In addition to the corpora, two lexicons were required, one for interjections the other for positive sentiment words. The lexicon for the interjections was taken from the searching various websites and compiling a list into a single text file. This consisted of two-hundred and seventy-two different interjections. The positive word lexicon was taken from Breen (2011). This was chosen over the more popular SentiWordsNet and WordNet lexicons because it only contained positive words, rather than a mixture of positive and negative. The lexicon contains a total of two-thousand and six positive sentiment words.

Due to the character limit used within Twitter, the tweets can often be incomplete and contain noise, such as hashtags and URL's, which are supplementary to the overall message (Jianqiang, 2015). Because of this, it was necessary to conduct pre-processing of the texts to increase the accuracy of the results whilst reducing processing times and the probability of errors. An important part of pre-processing of text data is the use of stopwords. These consist of common words which have little or no influence on the sentiment or emotion expressed. Examples of stopwords include; me, I, we, we're, for and they.

For this study, a number of pre-processing steps were taken. Because they are supplemental to the overall message, it was necessary to remove any emoticons, or emoji's, from the text in order for further processing to be carried out. Once the data had been converted the next step in pre-processing was to change the text into lowercase. By using lowercase all the text was then uniform and as the two lexicons were also all lowercased it allowed better string matches. After the conversion to lowercase the next steps were the removal of numbers and punctuations. The final step in the pre-processing was the removal of stopwords. For this task, a function was called which contained one-hundred and seventy-four of the most commonly used words. Additionally, appended to the stopwords were the terms sarcasm, sarcastic and trump. This was done as the tweets all contained either the hashtag #sarcasm or #trump and therefore it would have been presented as a feature in later stages.

Further pre-processing was conducted to remove a number of additional features which are often present within tweets. These included URL's, separate hashtags, retweets, extra whitespaces and user names. These were removed as they are considered surplus information and the removal can increase the processing speed.

For this study, all analysis was conducted using either the R programming language in R Studio or within Microsoft Excel. R and R Studio were chosen for this study as there exists a wide body of supportive literature, as well as the ease in which the twitteR package interacts with Twitter's API. Another consideration in the selection of R and R Studio was their open source licence, meaning it is freely available. Many other programming languages and development environments were considered, and some are deemed to be more powerful, but they are also prohibitively priced or require a much steeper learning curve. Microsoft's Excel was selected for similar reasons as it is widely available, supported and a powerful tool for data analysis. The selection of these two tools fitted with the overall aim of this study in being very rudimentary.

For the classification of text as being sarcastic using the two lexicons, of interjections and positive sentiments, a Boolean function was created to conduct an exact match of the words within the corpus to those within the lexicons. For each match found the total was calculated and stored within a data frame consisting of the tweet, total number of positive words matched and the total number of interjections matched. The data frame was then exported to a Microsoft Excel spreadsheet. Using the Excel function shown below, the tweets were classified as being either sarcastic or not.

```
=IF(AND(POS.MATCH>0,INT.MATCH>0),”Sarcastic”,”Not sarcastic”)
```

The Excel function looks at the results of the positive word matches and the interjection matches. If both are greater than zero, then the tweet is given the label of ‘Sarcastic’. If one or both matches are zero, then the tweet is labelled as ‘Not sarcastic’. The function was then copied across the range of the dataset to provide a label for each tweet. Once every tweet had been provided with a label another function was used to count the number of ‘Sarcastic’ labels. This function was as follows.

```
=COUNTIF(LABELS=”Sarcastic”)
```

This Excel function provides a total count for the tweets labelled as ‘Sarcastic’. It was then repeated for those tweets labelled as ‘Not sarcastic’. Using these two values a percentage of tweets labelled ‘Sarcastic’ was calculated. Additionally, the total number of words matched to each of the two lexicons was taken and a percentage calculated.

RESULTS

This task required the use of two lexicons and two corpora. The corpora were both taken from Twitter and both contained one-thousand and five-hundred records before pre-processing. One corpus was selected by using the hashtag #sarcasm. This was to be used as a test of the theory that positive sentiment and interjections could be used to classify short text as being sarcastic or not. The second corpus was selected by using the hashtag #trump.

The pre-processing reduced the corpus of sarcasm labelled texts to one-thousand, three-hundred and ninety-seven records. Pre-processing reduced the #trump corpus to eight-hundred and forty-four records. Each lexicon was compared to the corpora and the count of words matched was appended to the text. From these results a number of calculations were made, these included the number of positive words matched and the number of interjections matched. The result was a data frame consisting of three columns – text, positive words matched, interjections matched. The text was classified as being sarcastic if both positive words and interjections were present. Using the corpus with pre-marked sarcastic text, the results are as follows in Table 1.

	Total	Percentage
Positive words	590	42.2%
Interjections	1110	79.4%
Sarcasm classification	513	36.7%

Table 1: Results from the ‘bag-of-words’ classification of text labelled #sarcasm

As Table 1 shows, from a corpus of one-thousand, three-hundred and ninety-seven records, only five-hundred and ninety contained positive words. However, a much greater proportion contained interjections. The method utilised within this study was successful in classifying nearly thirty-seven percent of the text as being sarcastic. For the corpus labelled using the hashtag #trump, the results are shown in Table 2.

	Total	Percentage
Positive words	291	34.4%
Interjections	655	77.6%
Sarcasm classification	261	30.9%

Table 2: Results from the 'bag-of-words' classification of text labelled

Table 2 shows the results of the 'bag-of-words' classification for the corpus containing text labelled with the hashtag #trump. This corpus contained eight-hundred and forty-four records. Of these two-hundred and ninety-one records were found to contain positive words and six-hundred and fifty-five records contained interjections. Nearly thirty-one percent of the corpus was classified as containing sarcastic text.

DISCUSSION

The results for the two classification tasks were both consistent with each other. Both tasks found that interjections were the most common feature within the two corpora, with nearly eighty percent of the tweets containing this feature. Kreuz and Caucci (2007) believed that interjections were and important feature for the automatic identification of sarcasm. This study has found that, whilst it is a common feature within sarcastic tweets, interjections also feature heavily in non-sarcastic tweets.

Interestingly, only a small amount of the tweets contained words which were associated with positive sentiment. Within the corpus consisting of tweets labelled with #sarcasm, only forty-two percent of the posts were found to contain positive sentiment. For the other corpus, with tweets labelled #trump, only thirty-four percent of the tweets contained positive sentiment. This would suggest that the overall sentiment might not be such an important feature within sarcasm detection.

Both corpora were shown to contain, approximately, thirty percent of sarcastic tweets. As Joshi et al (2016) state, data skews are a problem within sarcasm detection. For the control classification, where a much higher result was expected as the tweets were labelled as being sarcastic by the author, manual annotation found six-hundred and five tweets were easily identifiable as being sarcastic. The remainder required additional context in the form of images or specialist knowledge. As context was not accounted for within this study, if these tweets are discounted then the result of the classification would be closer to eighty-five percent. Accounting for data skew within the corpus with the tweets labelled #trump, manual annotation found three-hundred and nineteen tweets were sarcastic. This would mean the method for classification would have successfully identified eighty-one percent of the sarcastic content.

The results indicate that positive sentiment and interjections are viable features for classifying sarcasm. However, at best, they are only accurate eighty percent. This would mean that a significant percentage of sarcastic content is misclassified as being not sarcastic. It would be interesting to establish if this could be accounted for with the use of negative sentiment.

Additionally, with the data skew found within the control corpus where each tweet was labelled as being sarcastic, finding a method which accounts for the additional context is important.

CONCLUSIONS

This study sought to establish whether the presence of both positive sentiment and interjections in a short text tweet is sufficient to be able to classify the text as containing sarcasm. The results showed that approximately thirty percent of each corpus was identifiable as being sarcastic using these features. This would indicate that they were unsuccessful. However, when accounting for data skews following a manual annotation of each corpus, the result was much greater and for both corpora the features achieved over eighty percent successful classification. These results would indicate that the aim of using such a rudimentary method for natural language processing is capable of achieving significant results. However, more complex machine learning based methods may have produced greater levels of accuracy.

The methods employed in this study, whilst not unique and are frequently used for learning sentiment analysis, are very basic. Many previous studies have used complex feature identification and machine learning algorithms to aid in the classification of sarcasm. This study has found that a much more rudimentary method is able to produce significant results with considerably less development time.

There are a number of limitations within this study. The classification task utilised two features found in previous studies for the detection of sarcasm in text. One of these features being positive sentiment. Sarcasm has been defined as being a method for an author to disguise their meaning by using an opposing sentiment. By focusing only on the positive sentiment, which would suggest a negative feeling, those tweets which contained negative sentiment, and therefore positive feeling, were ignored. Additionally, the use of interjections is not unique to sarcastic texts and many tweets may contain them where an author wishes to enhance the expressed sentiment. However, for the two corpora collected in this study this did not appear to be the case.

Overall this study has shown that basic methods of sarcasm detection produce mixed results. When employing a rudimentary feature identification for the creation of a sarcasm specific lexicon, this study found that the method did not produce viable outcome. However, the basic method of Boolean matches of two lexicons did show some success across two different corpora. Further research is needed into the nuances of sarcastic text and may require a more multi-dimensional approach which is able to account for context as well sentiment and interjections.

REFERENCES

Breen, J. (2011), “Twitter sentiment analysis tutorial,” available at: https://github.com/jeffreybreen/twitter-sentiment-analysis-tutorial-201107/blob/master/data/opinion-lexicon-English/positive-words.txt [accessed 03 March 2017].

Buschmeier, K. Cimiano, P. & Klinger, R. (2014) An impact analysis of features in a classification approach to irony detection in product reviews, in 5th Workshop on

Computational Approaches to Subjectivity, Sentiment and Social Media Analysis in Baltimore, USA – June 27, 2014, pp 42-49.

Davidov, D. Tsur, O. & Rappoport, A. (2010) Semi-supervised recognition of sarcastic sentences in Twitter and Amazon, in CoNLL '10 Proceedings of the Fourteenth Conference on Computational Natural Language Learning in Uppsala, Sweden — July 15 - 16, 2010, pp. 107-116.

Fernández-Gavilanes, M. Álvarez-López, T. Juncal-Martinez, J. Costa-Montenegro, E. & González-Castano, E. (2016), Unsupervised method for sentiment analysis in online texts, Expert Systems with Applications, Vol. 61, pp 57-75.

Filatova, E. (2012) Irony and sarcasm: corpus generation and analysis using crowdsourcing. Proceedings of the Eighth International Conference on Language Resources and Evaluation, LREC 2012 in Istanbul inTurkey, May 23-25, 2012. European Language Resources Association (ELRA) 2012, pp. 392-398.

Filik, R. Turcan, A. Thompson, D. Harvey, N. Davies, H. & Turner, A. (2016) Sarcasm and emotions: comprehension and emotional impact. Journal of Experimental Psychology. Vol 69 No. 11, pp 2130-2146.

González-Ibánez, R. Muresan, S. & Wacholder, N. (2010) Identifying sarcasm in Twitter: a closer look. Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies in Portland, Oregon — June 19 - 24, 2011, pp. 581-586.

Joshi, A. Bhattacharyya, P. & Carman, M. (2016), "Automatic sarcasm detection: a survey," available at: https://arxiv.org/pdf/1602.03426.pdf [accessed 10 March 2017].

Kovaz, D. Kreuz, R. & Riordan, M. (2013) Distinguishing sarcasm from literal language: evidence from books and blogging. Discourse Processes. Vol 50 No. 8, pp. 598-615.

Kreuz, R. & Caucci, G. (2007) Lexical influences on the perception of sarcasm, in Proceedings of the Workshop on Computational Approaches to Figurative Language, April 27, 2007, Association for Computational Linguistics, Lisbon, Portugal, pp. 1-4.

Liebrecht, C. Kunneman, F. & van den Bosch, A. (2013) The perfect solution for detecting sarcasm in tweets #not, in Proceedings of the 4th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, June 14, 2013, Association for Computational Linguistics, Atlanta Georgia, pp.29-37.

Lui, P. Chen, W. Ou, G. Wang, T. Yang, D. & Lei, K. (2014) Sarcasm detection in social media based on imbalanced classification, in The 15th International Conference on Web-Age Information Management, June 16-18, 2014, Springer International Publishing, Switzerland, pp 459-471.

Lukin, S. & Walker, M. (2013) Really? Well. Apparently bootstrapping improves the performance of sarcasm and nastiness classifiers for online dialogue, in Proceedings of the Workshop on Language Analysis in Social Media, North American Association for Computational Linguistics: Human Language Technologies, June 13-15, 2013, Association for Computational Linguistics, Atlanta Georgia, pp 30-40.

Maynard, D. & Greenwood, M. (2014) Who cares about sarcastic tweets? Investigating the impact of sarcasm on sentiment analysis. available at: http://www.lrec-conf.org/proceedings/lrec2014/pdf/67_Paper.pdf, [accessed 10 March 2017].

Pang, B. Lillian, L. & Vaithyanthan. (2002) Thumbs up? Sentiment classification using machine learning techniques, in Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing, July 6, 2002, Association for Computational Linguistics, Stroudsburg, Pennsylvania, pp. 79-86.

Reyes, A. & Rosso, P. (2014) On the difficulty of automatically detecting irony: beyond a simple case of negation. Knowledge and Information Systems, Vol. 40, No. 3, pp. 595-614.

ISSN 2309-0405

Reyes, A. Rosso, P. & Buscaldi, D. (2012) From human recognition to irony detection: the figurative language of social media. Data and Knowledge Engineering, Vol. 74, pp. 1-12.

Riloff, E. Qadir, A. Surve, P. De Silva, L. Gilbert, N. & Huang, R. (2013) Sarcasm as a contrast between a positive sentiment and a negative situation, in Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing, October 18-21, 2013, Association for Computational Linguistics, Seattle, Washington, pp. 704-714.

Thompson, D. & Filik, R. (2016) Sarcasm in written communication: emoticons are effective markers of intention. Journal of Computer-Mediated Communication. Vol. 21, No. 2, pp. 105-120.

Tsur, O. Davidov, D. & Rappoport, A. (2010) 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, May 23-26, 2010, Association for the Advancement of Artificial Intelligence, Menlo Park, California, pp. 162-169.

Turney, P. (2002) Thumbs up or thumbs downs? Semantic orientation applied to unsupervised classification of reviews, in Proceedings of the 40th Annual Meeting on Association for Computational Linguistics, July 7-12, 2002, ACM, Philadelphia, Pennsylvania, pp. 417-424.