"Opinion Mining in Twitter – Sarcasm Detection"

Sana Parveen[1], Sachin N. Deshmukh[2]

[1,2]Department of Computer Science and Information Techonology
Dr.Babasaheb AmbedkarMarathwada University,Maharashtra,India

Abstract - *Sarcasm is largely used in social networks and micro blogging websites, where people mock or criticize in a way that makes it difficult even for humans to tell if what is said is what is meant. To recognize sarcastic statements can be very useful when it comes to improving automatic sentiment analysis the data collected from social networks. This work demonstrated the importance of detecting sarcastic tweets automatically, and also demonstrate how the accuracy of sentiment analysis can be enhanced knowing which tweets are sarcastic and which are not. In this work we propose a method to detect sarcasm in Twitter that makes use of the different components of the tweet. Work proposes four categories of features that cover different types of sarcasm we defined, and that will be used to classify tweets into sarcastic and non-sarcastic. To evaluate the performances of our work study the importance of each of the proposed sets of features and evaluate its added value to the classification.*

Key Words: Opinion Mining, Sentiment Analysis, Sarcasm, POS Tagging.

1. INTRODUCTION

Sarcasm is a tongue of which the user speaks of something the complete opposite of what the user means. It often has the best comedic value. It is the use of irony to mock or convey contempt.

Ex. 1:
"I'm okay. Don't mind the gaping wound and the sword protruding from my back. I'm fine. Feel like a million bucks, dammit."

Ex. 2:
"Is your car stuck in the mud?"
"No, no, of course not, I'm only practicing how to spray mud using my tires".

In sarcastic manner the author or speaker usually speaks opposite of what he intend to say, it is highly dependent on the speakers intensity and speech patterns can detect sarcasm by the use of machine learning techniques. Sarcasm is the act of saying one thing while meaning the opposite. It is mostly a verbal device, with intention of putting someone down. For instance, if you say, "Yeah, he's a real mental giant" while rolling your eyes, you've just engaged in sarcasm. Though always mocking, sarcasm ranges from affectionate ribbing to deliberate humiliation. Sarcasm can be obvious, as in the example above, or it can be subtle or deadpan. Most people know someone who makes sarcastic remarks with a straight face, leaving his audience wondering if he meant what he said. That's because, on a literal level, the sarcastic remark could be true. For instance, if you say, "She's really beautiful," you could mean it. The tone and accompanying gestures are what let others know you are being sarcastic.

To analyze a sentence to detect sarcasm, context must be taken into account, as well as the tone of stressed syllables: English speakers tend to exaggerate tone when using sarcasm. Unfortunately, tone is not indicated in written English. We can't count all the times I've read conversations on the internet where the lack of tone in writing has caused sarcastic people to be mistaken as serious. Some people call it Poe's law. Several solutions have been proposed to resolve this, most involve introducing a new piece of punctuation, the "sarcasm/irony mark" which usually appears as a backwards question mark, or squiggly exclamation mark. Others, working inside the system as opposed to changing punctuation all together, use other punctuation enclosed in brackets to denote sarcasm ([?] or [!]), or add a fake HTML tag, </sarcasm>.

Detecting sarcasm is very important task in corporate and personnel word as if one fails to detect sarcasm in front of public users; it would ruin the image of product or company and person replying to the sarcastic comment. Mostly sarcasm has positive comments while user means negative feedback or the author shows positive attitude to show his negative opinion about the topic. Due to high data volume and speed of data generation, we need to automate the process of sarcasm detection and sentiment analysis.

Sarcastic statements are sort of a true lie. You're saying something you don't literally mean, and the communication works as intended only if your listener gets that you're insincere. Sarcasm has a two-faced quality: it's both funny and mean. This dual nature has led to contradictory theories on why we use it.

What makes task of detecting sarcasm hard is that even humans find it hard to understand them sometimes without prior knowledge of the topic. Sarcasm is also very closer to lie in some context, making it more problematic and hard task. As user or author writes exactly opposite of what he means, this is similar in lying. Sarcasm is widely used in twitter and other social networking websites, micro blogging is platform for sarcasm and twitter also has dedicator users for sarcasm. Sarcasm is having intense words in its structure giving more pressure on the use of intense words that makes human understand sarcasm. Making it even worse to detect sarcasm, Data structure of twitter [i] is more informal immature with an evolving vocabulary of slang words and abbreviations and [ii] has a limit of 140 characters per tweet which provides fewer word-level cues thus adding more ambiguity.[1]

2. RELATED WORK

Automatic detection of sarcasm could be a comparatively new, less researched topic and is deemed a troublesome problem (Pang and Lee). Whereas works on automatic detection of sarcasm in speech (Tepperman et al.) utilizes speech, spectral and contextual options, sarcasm detection in text has relied on characteristic text patterns (Davidovet al.) [2] and lexical features (Gonz´alez-Ib´a˜nez et al; Kreuz) [3]. Current works on sarcasm detection have heavily focused on sarcasm‟s linguistic aspects and utilized primarily, the content of the tweet. Liebrecht et al. introduce a sarcasm detection system for tweets, messages on the micro blogging service offered by Twitter.

In micro-blogging sites like Twitter, tweets are typically expressly marked with the #sarcasm hashtag to point that it's satirical. Analysis has shown that sarcasm is usually signaled by exaggeration, using intensifiers and exclamations. In distinction to the present, non-hyperbolic sarcastic messages typically have a particular marker. not like an easy negation, a sarcastic message conveys a negative opinion using solely positive words or intense positive words. In step with Gibbs and Izett, sarcasm divides its addressees into 2 groups; first: individuals who perceive sarcasm and a group of individuals who don't perceive sarcasm. On Twitter, the senders use the hashtag so as to confirm that the addressees sight the sarcasm in their text.

Target a brand new approach to sentiment analysis by using "word senses" as "semantic features" for sentiment classification. In his paper, he used WordNet 2.1 (Fellbaum) because the sense repository every word is mapped to a synset based on its sense [4].

Pang et al. in their paper analyses the performance of unigram as features. The results showed that unigram presence taken as feature seems to be the most effective. This work contains n-gram as options so as to capture the context. However the paper's experimental results showed that bigram as feature failed to improve the performance of the sentiment classifier to any extent further. So, unigram features are most popular over n-gram features.

Liebrecht et al. developed and tested a system that detects sarcastic tweets in a realistic sample of 3.3 million Dutch tweets posted on a single day, trained on a set of nearly 78 thousand tweets, harvested over time, marked by the hashmark #sarcasme by the senders.[5]

Zang et al. constructed a deep neural network model for tweet sarcasm detection [6]. Compared with traditional models with manual discrete features, the neural network model has two main advantages. First, it is free from manual feature engineering and external resources such as POS taggers and sentiment lexicons. Second, it leverages distributed embedding inputs and recurrent neural networks to induce semantic features. The neural network model gave improved results over a state-of-the-art discrete model. In addition, we found that under the neural setting, contextual tweeter features are same effective with both sarcasm detection and with discrete models [7].

Table -1: Overview of Work

Overview of work	**Authors & Year**
Bouazizi et al. (2015)	Extracted features from the tweets and used machine learning to run the classification.[7][8]
Riloff et al. (2013)	Lexicon-based approach contrasting positive sentiment and negative situation [9]
Liebrecht.et.al. (2013)	Unigram, bigram and trigram features used to train a Balanced Winnow classifier[5]
Reyes et al. (2012)	Ambiguity, polarity, emotional cues etc., to train decision trees [10]
Zang et al. (2016)	Deep Neural networks with semantic features for sarcasm detection from tweets [6]

3. METHODOLOGY

In this proposed work, we have created two datasets i.e. before adding sarcasm tweets into training data and after adding sarcasm tweets into training data as described in Table 2 and Table 4 respectively.

Table-1: Structure of Dataset before adding Sarcasm Tweets

Topic	No. of Tweets	
	Training	Testing
General	1200	125
Sport	1200	125
Phone Reviews	1200	125
Movie Reviews	1200	125
Electronic Products Reviews	1200	125
Politics	1200	125
Total	7200	750

Given Topic contain 1200 tweets as a Training data and 125 tweets as a Testing data in which 25 tweets are sarcastic towards their topic and remaining 100 tweets are non-sarcastic. For total 7200 tweets as a Training data and 750 tweets for Testing data, on which we have performed classifications like Naive Bayes, Maximum Entropy and Support Vector Machine algorithms.

We preprocessed review data for non ASCII characters, special characters blank lines and spaces, removed punctuations and remarks. We reduced three or more consecutive characters to two. We then lemmatized out review data. After preprocessing all characters were converted to small letters and numbers were eliminated.
Next step comes POS tagging of words, it was done using penn tree bank to tag each word in review sentences with the part of speech associated with it. Training data was pure and genuine reviews with no sarcasm. We mixed and shuffled #SARCASM tweets with all reviews and performed classification.

We have six categories of reviews having 1325 tweets in each category and sarcasm tweets mixed with them for better detection of sarcasm and helping improve sentiment analysis of our algorithms with sentiment pre labeled to them for training and evaluation of model. Work performed sentiment analysis on the data. In developing our model our first experiment was conducted on simple reviews with no sentences marked as sarcastic, in second experiment we added 2800 tweets to training dataset and performed sentiment analysis.

We trained our classifiers with 4 types of features extracted from the training data which were sentiment-related features, punctuation-related features, syntactic features and pattern features etc. These features were found to be different in sarcastic tweets thus classifiers knew the properties of the sarcastic reviews and classified them negative sentiment.

It was observed that before adding sarcasm tweets to data as they are negative sentiment data, accuracy of all the classifiers was increased.

4. EXPERIMENTAL WORK AND RESULTS

In the results we find that, the 25 sarcastic tweets in testing data from each topic wasn't detect correctly. So it affects the performance of classifiers. So to improve this, we add 2800 sarcastic tweets in training data. So now, we have 10000 tweets as a training data and 750 tweets as a testing data that shown in Table 4 .Table 3 shows the performance of NB, ME and SVM classifiers.

Table -3: Result of classifiers before adding Sarcasm Tweets

Algorithms	No. of Tweets	Result	
	Training	Testing	
Naive Bayes	1200	125	60.66 %
Support vector Machine	1200	125	68.93 %
Maximum Entropy	1200	125	74.53 %

Table-2: Structure of Dataset after adding Sarcasm Tweets

Topic	No. of Tweets	Testing
	Training	
General	1200	125
Sport	1200	125
Phone Reviews	1200	125
Movie Reviews	1200	125
Electronic Products Reviews	1200	125
Politics	1200	125
#SARCASM Tweets	2800	-
Total	10000	750

Once completed adding sarcasm tweets into training data, we have extracted the features i.e. sentiment-related features, punctuation-related features, syntactic features and pattern features from training data performed detection of sarcasm. After extracting features, we perform machine learning algorithms on it. Table 5 shows the results of classifiers after adding sarcasm tweets into training data and extracting features rom training data.

Table-3: Result of classifiers after adding Sarcasm Tweets

Algorithms	No. of Tweets	Results	
	Training	Testing	
Naive Bayes	1200	125	67.06 %
Support vector Machine	1200	125	75.33 %
Maximum Entropy	1200	125	82.13 %

Classification is conducted using Naive Bayes, SVM, and Maximum Entropy algorithms. Table 6 and Figure 1 show the accuracy of sentiment classification before and after taking sarcasm-related features into consideration. The results show a noticeable enhancement after taking the sarcasm into consideration. Albeit the low number of sarcastic tweets in our test set (i.e., less than 5%), our approach helped enhance the results. In other words, many of the tweets, previously classified wrongly are now well classified.

Table-4: Accuracy of Sentiment Classification

Classifiers	Result	After
	Before	
Naive Bayes	60.66 %	67.06 %
Support vector Machine	68.93 %	75.33 %
Maximum Entropy	74.53 %	82.13 %

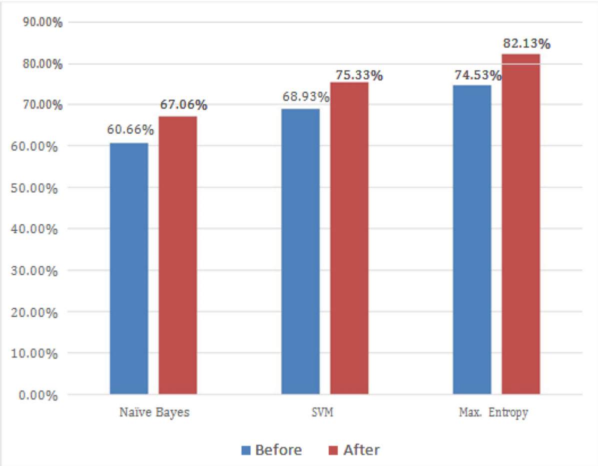

Chart-1: Accuracy of Sentiment Classification

5.CONCLUSION

Opinion mining and sentiment analysis refer to the identification and the aggregation of attitudes or opinions expressed by internet users towards a specific topic. Analyzing the sentiment of tweets gives an interesting insight into the opinions of the public in relation to a certain event. However, due to the limitation in terms of characters (i.e. 140 characters per tweet) and the use of informal language, the state-of-the-art approaches of sentiment analysis present lower performances in Twitter than that when they are applied on longer texts. Moreover, presence of sarcasm makes the task even more challenging. Sarcasm is when a person conveys implicit information, usually the opposite of what is said, within the message he transmits. Future work will focus on the polarity classification of scalable topic-level streaming feeds, with classification of a streaming feeds' sentiment towards a given topic (and not just a keyword). The next step can be defined as; trend detection relating to a topic on a set of streaming feeds, to determine the polarity of the target topics. Also, determining the degree of polarity can be used to show the sentiment strength (such as strongly positive/negative or weakly positive/negative or neutral).

REFERENCES

[1] S. Chen, B. Mulgrew, and P. M. Grant, "A clustering technique for digital communications channel equalization using radial basis function networks," IEEE Trans. on Neural Networks, vol. 4, pp. 570-578, July 1993.

[2] J. U. Duncombe, "Infrared navigation—Part I: An assessment of feasibility," IEEE Trans. Electron Devices, vol. ED-11, pp. 34-39, Jan. 1959.

[3] C. Y. Lin, M. Wu, J. A. Bloom, I. J. Cox, and M. Miller, "Rotation, scale, and translation resilient public watermarking for images," IEEE Trans. Image Process., vol. 10, no. 5, pp. 767-782, May 2001.

[4] M. Bouazizi, T. Ohtsuki, "Pattern-Based Approach for Sarcasm Detection on Twitter".

[5] E. Fersini, F. A. Pozzi, and E. Messina, ''Detecting irony and sarcasm in microblogs: The role of expressive signals and ensemble classifiers,''. in Proc. IEEE Data Sci. Adv. Anal. (DSAA), Oct. 2015, pp. 1–8.

[6] Z. Wang, Z. Wu, R. Wang, and Y. Ren, ''Twitter sarcasm detection exploiting a context-based model,'' in Proc. Web Inf. Syst. Eng. (WISE), Nov. 2015, pp. 77–91.

[7] DD. Bamman and N. A. Smith, ''Contextualized sarcasm detection on Twitter,'' in Proc. AAAI Int. Conf. Web Soc. Media (ICWSM), May 2015,pp. 574–577.

[8] J. D. Campbell and A. N. Katz, ''Are there necessary conditions forinducing a sense of sarcastic irony?'' Discourse Process., vol. 49, no. 6,pp. 459–480, Aug. 2012.

[9] E. Riloff, A. Qadir, P. Surve, L. De Silva, N. Gilbert, and R. Huang,''Sarcasm as contrast between a positive sentiment and negative situation,'' in Proc. Conf. Empirical Methods Natural Lang. Process., Oct. 2013,pp. 704–714

[10] A. Reyes, P. Rosso, and T. Veale, ''A multidimensional approachfor detecting irony in Twitter,'' Lang. Resour. Eval., vol. 47, no. 1,pp. 239–268, Mar. 2013.