Opinion Mining in Twitter: How to make use of Sarcasm to Enhance Sentiment Analysis: A Review

Sana Parveen, Avinash Surnar, Sunil Sonawane

***Abstract*—NLP is processing linguistic data to get its features or to better understand the human languages. It mainly involves Processing of big natural language corpora by computer to get a desired output as information from the corpora. In opinion mining, we determine mood of Author from the text about a particular topic or product. It is also called Sentiment analysis and its task involves building a system that categorizes the text data. The Free Dictionary1 defines sarcasm as a form of verbal irony that is intended to express contempt or ridicule. This paper states the work that is done identifying sarcasm in the twits to improve the opinion mining in twitter data.**

***Index Terms*—Feature extraction, Opinion Mining, Pattern classification, Sarcasm.**

I. INTRODUCTION

Sarcasm is a nuanced form of language where usually, the speaker explicitly states the opposite of what is implied. Its synonyms are derision, mockery, ridicule, satire, irony, scorn, sneering, scoffing, gibing, taunting; trenchancy, mordancy, causticity, mordacity. Although sarcasm is largely dependent on the tone of the speaker, computers can detect sarcastic sentences on the sentence basis using machine learning techniques. 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.

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]

II. RELATED WORK

Automatic detection of sarcasm could be a comparatively new, less researched topic and is deemed a troublesome problem (Pang and Lee, 2008). Whereas works on automatic detection of sarcasm in speech (Tepperman et al., 2006) utilizes speech, spectral and contextual options, sarcasm detection in text has relied on characteristic text patterns (Davidovet al., 2010) [2] and lexical features (Gonz´alez-Ib´a˜nez et al., 2011; 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. (2013) 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

***Manuscript received May, 2017*.**

Sana Parveen, *Department of Computer Science and Information Technology, Dr. BabasahebAmbedkarMarathwada University, Aurangabad, India, +9422631344*

Avinash Surnar, *Department of Computer Science and Information Technology, Dr. BabasahebAmbedkarMarathwada University, Aurangabad, India, +919921615326.*

Sunil Sonawane, *Department of Computer Science and Information Technology, Dr. BabasahebAmbedkarMarathwada University, Aurangabad, India, +918483927903.*

and Izett (2005), 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, 1998) because the sense repository every word is mapped to a synset based on its sense [4].

Pang et al. (2002) 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. We are able to have 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.(2013) 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. (2016) constructed a deep neural network model for tweet sarcasm detection. 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 [6].

Table: Overview of Work

Authors & Year	**Overview of work**
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]

III. METHODOLOGY

Social media is source of vast form of information and data is continuously posted on social media. However, because of limit for words to be posted, on some platforms data tends to be small length. However, tweets have been used more by researchers for text classification and sentiment analysis. This might be as a result of convenience of the Twitter API and recognition of twitter as a medium.

A. Data

There are three ways to create dataset of sarcastic tweets to detect sarcasm in tweets.

1. Manual annotation
2. Hashtag based and
3. Context based

Manual Annotation of tweets is first of the techniques to create tweeter dataset for sarcasm detection..Riloff et al. [2013]annoted tweets whether they are sarcastic or not. Maynard and Greenwood [2014] used their dataset to study impact of sarcasm on tweeter.

Hashtag based dataset is second technique, where hash tag are used to specify if tweets are sarcastic. Many researchers use hashtag based data to categorize the tweeter data for sarcasm and other purposes. Limitation of hashtag based data is:

(a) Authors of tweets provide hastags only they know if tweets are sarcastic or not,

(b) The approach allows creation of large-scale datasets.

For creating this type of dataset, tweets are labeled with a new type of hashtag called "sarcastic".

Davidov et al. [2010] used dataset, where tweets were labeled with hashtags such as #sarcasm, #sarcasticetc.

Gonzalez-Ib ´ anez et al. [2011] also used hashtag-based supervision for tweeter data. They only considered tweets with hashtag at end of tweet like suffix. Reyes et al. [2012] use similar approach.

Liebrecht et al. [2013] used '#not' hashtagto consider whether there tweet are sarcastic and label them.

Barbieri et al. [2014b] create a dataset using hashtag-based supervision based on hashtags indicated by multiple labels: Other works using this approach have also been reported

[Barbieri et al. 2014a; Joshi et al. 2015; Bharti et al. 2015; Bouazizi and Ohtsuki 2015a; Abercrombie and Hovy 2016]. Requirement of quality control in distant supervision makes sarcasm detection hard.

To keep quality, Bamman and Smith [2015] labeled tweets as positive if they have hashtag sarcasm and negative if they don't.

Fersini et al. [2015] present a dataset of 8K tweets where the initial label is based on the hashtag. To ensure quality, these tweets are labeled manually [11]

In order to detectsarcasm tweets, supplementary datasets have also been created for sarcasm detection. Khattri et al. [2015] use a supplementary set of complete twitter timeline to establish context for a given dataset of tweets.

[Rajadesingan et al. 2015] used a twitter data, labeled by hashtag-based criteria along with a context related data of 80 tweets per author [12].

B. Features

Now we are going to see the features that are considered while detecting a sarcastic tweet, most techniques have used bag-of-words as features.

Though most work is done in unigram and n-gram features some work is done on pragmatic, lexicon based and pattern based features of the tweets. Tsur et al. [2010] design pattern-based features that indicate presence of discriminative patterns as extracted from a large sarcasm-labeled corpus. To allow generalized patterns to be spotted by the classifiers, these pattern-based features take real values based on three situations: exact match, partial overlap and no match.

Gonzalez- ´ Ibanez et al. [2011] use sentiment lexicon-based features. In addition, pragmatic features like emoticons and user mentions are also used. Reyes et al. [2012] introduce features related to ambiguity, unexpectedness, emotional scenario, etc. Ambiguity features cover structural, morpho-syntactic, semantic ambiguity, while unexpectedness features measure semantic relatedness.

Riloff et al. [2013] used positive verbs and negative situation phrases for classifieras pattern based features.

Bigrams and trigrams as featuresLiebrecht et al. [2013] skip-gram and character n-gram-based features Reyes et al. [2013].

Maynard and Greenwood [2014] include seven sets of features. Some of these are maximum/minimum/gap of intensity of adjectives and adverbs, max/min/average number of synonyms and synsets for words in the target text, etc. Apart from a subset of these [13]

Barbieri et al. [2014a] use frequency and rarity of words as indicators [14]. Buschmeier et al. [2014b] incorporate ellipsis, hyperbole and imbalance in their set of features [15].

Joshi et al. [2015] use features corresponding to the linguistic theory of incongruity. The features are classified into two sets: implicit and explicit incongruitybased features. Hernandez-Far ´ıas et al. [2015] present features that measure semantic relatedness between words using Wordnet-based similarity [16].

C. Machine Learning Techniques

SVM has been used most of the times for detection of sarcasm [Joshi et al. 2015; Tepperman et al. 2006; Kreuz and Caucci 2007; Tsur et al. 2010; Davidov et al. 2010] (or SVM-Perf as in the case of Joshi et al. [2016b]).

Gonzalez-Ibanezet al. [2011] use SVM with SMO ´ and logistic regression.Chi-squared test is used to identify discriminating features.

Reyes and Rosso [2012] use Naive Bayes and SVM. They also show Jaccard similarity between labels and the features.

Riloff et al. [2013] compare rule-based techniques with a SVM-based classifier. Liebrecht et al. [2013] use balanced winnow algorithm in order to determine high-ranking features.

Reyes et al. [2013] use Naive Bayes and decision trees for multiple pairs of labels among irony, humor, politics and education.

Bamman and Smith [2015] use binary logistic regression. Wang et al. [2015] use SVMHMM in order to incorporate sequence nature of output labels in a conversation. Liu et al. [2014] compare several classification approaches including bagging, boosting, etc. and show results on five datasets.

On the contrary, Joshi et al. [2016a] experimentally validate that for conversational data, sequence labeling algorithms perform better than classification algorithms. They use SVM-HMM and SEARN as the sequence labeling algorithms.

IV. Result and Discussion

Trends observed in sarcasm detection research as seen in the figure, there have been four key milestones. Following fundamental studies, supervised/semi-supervised sarcasm classification approaches were explored. These approaches focused on using specific patterns or novel features. Then, as twitter emerged as a viable source of data, hash tag-based supervision became popular. Recently, using context beyond the text to be classified has become popular. two of these trends: (a) discovery of sarcastic patterns, and use of these patterns as features, and (b) use of contextual information i.e., information beyond the target text for sarcasm detection are most popular.

V. Conclusion

In Sarcasm detection research has grown significantly in the past few years, necessitating a look-back at the overall picture that these individual works have led to. This paper surveys approaches for automatic sarcasm detection. We observed three milestones in the history of sarcasm detection research: semi-supervised pattern extraction to identify implicit sentiment, use of hashtags-based supervision, and use of context beyond target text.

REFERENCES

[1] http://www.cs.cornell.edu/home/llee/opinion-mining-sentiment-analysis-survey.html

[2] 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).

[3] Gonzalez-Ibanez, R., Muresan, S., &Wacholder, N. “Identifying sarcasm in Twitter: a closer look”, In Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: short papers-Volume 2 (pp. 581-586). Association for Computational Linguistics

[4] Joseph Tepperman, D. R. “Yeah right: sarcasm recognition for spoken dialogue systems”, In Proceedings of INTERSPEECH

[5] C.C Liebrecht, F. K. “The perfect solution for detecting sarcasm in tweets”, Proceedings of the 4th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, pages 29–37

[6] Meishan Zhang, Yue Zhang, and Duy-Tin Vo. “Gated neural networks for targeted sentiment analysis”, isnthirtieth sAAAI Conference on Artificial Intelligence.

[7] MondherBouazizi and TomoakiOhtsuki “Opinion Mining in Twitter How to Make Use of Sarcasm to Enhance Sentiment Analysis”, In Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015. ACM, 1594–1597.

[8] MondherBouazizi and TomoakiOhtsuki “Sarcasm Detection in Twitter: All Your Products Are Incredibly Amazing!!!”-Are They Really?” In 2015 IEEE Global Communications Conference (GLOBECOM). IEEE, 1–6.

[9] Ellen Riloff, AshequlQadir, PrafullaSurve, Lalindra De Silva, Nathan Gilbert, and Ruihong Huang “Sarcasm as Contrast between a Positive Sentiment and Negative Situation”, In conference EMNLP. 704–714.

[10] Antonio Reyes and Paolo Rosso “Making objective decisions from subjective data: Detecting irony in customer reviews”, In Decision Support Systems 53, 4 (2012), 754–760.

[11] David Bamman and Noah A Smith, “Contextualized Sarcasm Detection on Twitter”, In Ninth International AAAI Conference on Web and Social Media.

[12] Rajadesingan, Ashwin, Reza Zafarani, and Huan Liu. "Sarcasm detection on Twitter: A behavioral modeling approach." Proceedings of the Eighth ACM International Conference on Web Search and Data Mining. ACM, 2015.

[13] Diana Maynard and Mark A Greenwood. 2014, “Who cares about sarcastic tweets? investigating the impact of sarcasm on sentiment analysis”, In Proceedings journal of LREC

[14] Francesco Barbieri, Francesco Ronzano, and HoracioSaggion. 2014a. “Italian irony detection in twitter: a first approach”, In The First Italian Conference on Computational Linguistics CLiC-it 2014 & the Fourth International Workshop EVALITA, pages 28–32.

[15] Francesco Barbieri, HoracioSaggion, and Francesco Ronzano. 2014b. “Modelling sarcasm in twitter, a novel approach”, ACL 2014, page 50.

[16] Aditya Joshi, Vinita Sharma, and Pushpak Bhattacharyya. “Harnessing context incongruity for sarcasm detection”, In Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing, volume 2, pages 757–762.