



## Issues and Opinions

## Sarcasm detection in microblogs using Naïve Bayes and fuzzy clustering



Shubhadeep Mukherjee\*, Pradip Kumar Bala

Indian Institute of Management, Ranchi, Information Systems Area, India

## ARTICLE INFO

## Article history:

Received 18 January 2016

Received in revised form

21 September 2016

Accepted 31 October 2016

Available online 4 November 2016

## ABSTRACT

Sarcasm detection of online text is a task of growing importance in the globalized world. Large corporations are interested in knowing how consumers perceive the various products launched by the companies based on analysis of microblogs, such as - Twitter, about their products. These reviews/comments/posts are under the constant threat of being classified in the wrong category due to use of sarcasm in sentences. Automatic detection of sarcasm in microblogs, such as - Twitter, is a difficult task. It requires a system that can use some knowledge to interpret the linguistic styles of authors. In this work, we try to provide this knowledge to the system by considering different sets of features which are relatively independent of the text, namely - function words and part of speech n-grams. We test a range of different feature sets using the Naïve Bayes and fuzzy clustering algorithms. Our results show that the sarcasm detection task benefits from the inclusion of features which capture authorial style of the microblog authors. We achieve an accuracy of approximately 65% which is on the higher side of the sarcasm detection literature.

© 2016 Elsevier Ltd. All rights reserved.

## 1. Introduction

Macmillan English dictionary defines sarcasm as the activity of saying or writing the opposite of what one means or of saying in a way intended to make someone else feel stupid or show them that one is angry [1]. With sophistication of language, use of sarcasm in verbal and written text has become quite the norm. However, automatic detection of sarcasm is still in its infancy. The ambiguous nature of sarcasm makes it difficult even for humans to detect it in sentences.

Despite the difficulties, the huge benefit of detecting sarcasm has been recognized in many computer interaction based applications, such as, review summarization, dialogue systems and review ranking systems [2]. From a business perspective, detecting sarcasm can be crucial in understanding product reviews, movie popularity and social opinions, all of which suffer a high threat of being recognized in the wrong category, when reviews and opinions are laden with sarcasm. Sarcasm, being a special type of communication, where the implicit meaning differs from the explicit one, cannot be effectively identified only by conventional data mining techniques.

It goes without saying, that sarcasm detection from unstructured text data is a relevant and challenging problem. It is without any visual or vocal aids that assist humans in understanding sarcasm. One of the major issues in sarcasm detection is the absence of naturally occurring expressions that can be used for training purposes [2]. In case of microblogs, such as, Twitter, messages can be annotated with hashtags that are an indication of the sentiment being expressed in tweets. These hashtags are reliable indicators of the emotion being expressed by the tweets, as the author explicitly conveys the emotion of the sentence through them (e.g. - #happy, #joy, #sad). We utilized this behavior to formulate two types of hashtags (#sarcasm, #notsarcasm, similarly #sarcastic, #notsarcastic) for our dataset. We consider the sentences that end in #sarcasm or #sarcastic to be the gold standard for sarcastic sentences and the sentences that end with the hashtag #notsarcasm or #notsarcastic to be the gold standard for non-sarcastic sentences. We do supervised learning, using Naïve Bayes classifier, using a class of features to differentiate a sarcastic tweet from a non-sarcastic one. In addition to that, we also do unsupervised learning using fuzzy c – means clustering.

There have been studies in the past that have tried to capture sarcasm through content based linguistic features [3], however, these studies have majorly relied on words used, emoticons and the sentences in general to differentiate between a sarcastic and a non-sarcastic sentence. This approach has worked reasonably well for

\* Corresponding author. Indian Institute of Management, 5th Floor SuchnaBhawan, Meurs Road, Ranchi, Jharkhand, 834001, India.

E-mail address: [shubhadeep.mukherjee13fpm@iimranchi.ac.in](mailto:shubhadeep.mukherjee13fpm@iimranchi.ac.in) (S. Mukherjee).

the datasets used, however, the performance of the algorithms, dependent solely on content based linguistic features, is likely to go down when applied to other datasets, due to their high dependency on the words used. One could argue that the results obtained in such cases are not generalizable to a satisfactory extent. Hence, in our proposed algorithm, we have derived and applied features that are relatively independent of the dataset making them universally applicable for sarcasm detection.

The feature set we considered for the classification of tweets consist of content words, function words, part of speech tags, part of speech n-grams and their various combinations. As a first in the field, we have used topic as well as style based features to classify the tweets for sarcasm detection. We did not come across any work in the field which has tried to capture authorial style based features for sarcasm detection. Our method thus adds a new dimension to natural language processing based research on sarcasm detection.

An English sentence can be broadly said to consist of two types of words - function words and content words. Function words are words that have little or no significant meaning outside the premise of the sentence. On the other hand, content words are words that have meaning even outside the context of the sentence [4]. Example of function words are – the, and, he, not etc. The examples of content words are – school, dog, angry etc. If we consider an English sentence in its entirety, it would consist of these two categories of words.

We hypothesize that, sarcasm in a sentence is dependent on the content of the sentence as well as the authorial or writing style of the author. The authorial or writing style is best captured by the function words and the part of speech used in the sentence [5]. Koppel et al. (2006) states that categorization by topic is typically based on keywords that reflects a document's content whereas categorization by author style uses precisely those features that are independent of context. Authorial style based classification has been applied successfully in gender classification of regular text [6,7]. We propose that the content of the tweets as well as the authorial style both contribute to the extent of sarcasm present in the tweets. We have used features that are independent of the content of the text in conjunction with other topic or content based features. By content based features, we mean those features which are an integral part of the text and give the text its meaning. For example – if we consider a sentence “My house is near the cave,” then the content words are – “house”, “near” and “cave”. The rest - “is” and “the” are function words or writing style based features that vary from author to author.

Before delving further, we would like to discuss some of the various sarcastic tweets that we have come across in our tweet datasets:

- a. This day just keeps getting better and better..!!
- b. I'd like to thank Michele Obama for making the fruit snacks in the lunch room 90% tinier! Really changed my whole life with that one.
- c. I just love that people are abusive and can get away with it.
- d. Those are my two cents. But what do I know? #idonthaveamasters
- e Because all Middle Eastern people are Muslims, right? #ttrttt

All the above sentences convey sarcasm by using certain style. The first sentence apparently conveys disgust at the way the user's day went from bad to worse by saying the exact opposite of it. The second sentence is a mockery at the fight against Obesity project started by Michelle Obama the first lady of the United States of America. The third sentence is a rebuttal of the abusive behavior of people in general. The fourth sentence is used to lament the fact that opinions of educated people matter more. This is understood

by the use of hashtag that says that the author doesn't have a masters. The last sentence is a dig at the generalization of people belonging to a region of being of a certain religion as well. All the above sentences convey sarcasm. In this work, we have tested the feature types that capture the authorial style and the content across five Twitter datasets. Our experiments reveal, that a combination of authorial style and content words of the tweet perform better than the feature types solely based on content or writing style.

## 2. Literature review

We have divided the literature review section into two parts. In the first part, we do a generalized overview of the research conducted in understanding sarcasm and its use across areas. In the secondpart, we do a more specific review of the work done in automatic sarcasm detection which is the premise of our research.

### 2.1. Overview

Sarcasm is a form of speech act in which the speakers convey their message in an implicit way [2]. The implicitness of the statements makes it hard for humans to decide whether a statement is sarcastic or not. Sarcasm has been studied in-depth in linguistics, psychology and cognitive sciences [8–11]. Jorgenson et al. (1984) stated that sarcasm arises from figurative meaning as opposed to literal meaning. Clark and Gerrig (1984) proposed that sarcasm cancels the indirectly negated message by replacing it with the implicated one. Giora (1995) refuted the claims of the earlier researchers by stating sarcasm to be a mode of indirect negation which requires processing of both the negated and implicated messages. Later, Ivanko and Pexman (2003) studied the inherent complexity of sarcasm and its effect on sarcasm processing time.

The following table summarizes the major works in sarcasm: (see Table 1)

### 2.2. Automatic sarcasm detection of online text:

One of the remarkable work on sarcasm detection in the field of text mining has been done by Tsur et al. (2010). The authors used a semi supervised algorithm for sarcasm detection (**SASI**) in product reviews. It consisted of 2 stages: semi supervised pattern acquisition and sarcasm classification. They used Amazon review for books and products for the task. The pattern acquisition task in their work consisted of pattern extraction, selection and matching. Additionally, they also used punctuation based features for classification. Davidov et al. (2010) used semi supervised learning based on **SASI** to classify tweets and amazon product reviews. Gonzalez-Ibanez et al. (2011) used basic supervised learning techniques to classify tweets for sarcasm using hashtags (#sarcasm) as gold standard. The authors used both lexical and pragmatic features for the classification job. They also did a comparative study of the human performance with the machine learning algorithm on accuracy of sarcasm detection. The authors also identified most discriminating features using multiple classes to gain insights in the problem. Most recently, Justo et al. (2014) have used a range of different features, such as, unigrams to classify tweets for sarcasm.

The following table summarizes the work done in the field of automatic sarcasm detection.(see Table 2)

Psychological research on sarcasm has revealed that use of sarcasm is associated with socio-economic class and profession [18]. For instance, a comedian is much more likely to use sarcasm in his/her sentences than a school teacher. One could say that a person's profession, habits, social circle affect the way he/she thinks, talks and writes [19]. This effectively makes the author's writing style all the more critical than merely the meaningful content of

Download English Version:

<https://daneshyari.com/en/article/4941800>

Download Persian Version:

<https://daneshyari.com/article/4941800>

[Daneshyari.com](https://daneshyari.com)