

Critique 2: “Contextualized Sarcasm Detection on Twitter”

David Bamman and Noah A. Smith

While the paper’s proposal using contextual cues to detect sarcasm yields promising findings, the paper lacks an analysis on the nature of sarcasms in communication terms. This limitation causes several gaps in the data collection process and analysis of conclusion.

Sarcasm as a type of communication, has different forms, with different purposes in communication. Among different categorizations of sarcasm in literature, based on purpose, one classification is “sarcasm as wit (being funny), as whimper (showing angry feeling), and as evasion (avoiding giving clear answer)”^[1]. Given the set-up of Twitter, sarcasm as evasion would hardly exist, because people have the choice to avoid, not post at the first place. Meanwhile, since the audience of Twitter posts are the public. This is different from private conversation where sarcasm could target to a specific person. For example, “You are so smart!!” would occur in private conversation, but not in a post addressing the public. This dimension shifts the dynamics of sarcasm making the object of sarcasm less personal. Thus, it entails careful measure to apply a rule based on one kind of sarcasm and apply it generally to the whole genre. A solution is setting the boundaries, specifying the characteristics of one communication channel and audience, focusing on one type of sarcasm.

In data collection, the paper adopts the self-declaration method to label sarcasm. The specific method is using #sarcasm tag from twitter to mark positive samples. As the paper states, “the design does not capture ... the varieties of sarcasm expressed without an explicit hashtag”. Indeed, though ensuring the accuracy of positive sample this would omit other untagged sarcastic posts in positive sample. More importantly, the negative sample may include those omitted sarcastic posts which should have positive labels. Further, since using hashtag is a behavioral choice on Twitter, studying the labeled sarcastic posts has two folds, sarcastic posts, and labeling behaviors. To solve the problem, the paper could use a combination of human annotator and self-declaration method. For untagged posts, have one or several volunteers to judge and keep the agreed non-sarcastic posts in the negative sample. Another way to apply self-declaration method is to establish contact with the authors and reconfirm their intent.

Coming to analysis of model finding, little discussion is on making sense of the contextual cues, how they relate to sarcasm detection on Twitter. For example, it’s worth exploring why author historical salient terms (81.2%) is a top ranked feature, how this feature relates to sarcasm. A large portion of discussions are on the distinction between tagged sarcasm and sarcasm between friends. It makes sense if it is pointing out the limitation of the paper; otherwise, this departs from stated goal of sarcasm detection, in which case comparing objects should be sarcastic posts and non-sarcastic posts.

In conclusion, presenting the difference between tagged sarcasm and sarcasm among friends as a limitation in conclusion is too late. Two assumptions are worth examining before defining the research question, how contextual cues could improve detection sarcasm on Twitter: first, is sarcasm on social media the same as it is in traditional face-to-face communication? Second, are there different types of sarcasm on Twitter, besides the form of tagged sarcasm?

1. M. Bouazizi and T. Otsuki Ohtsuki, "A Pattern-Based Approach for Sarcasm Detection on Twitter," in *IEEE Access*, vol. 4, pp. 5477-5488, 2016, doi: 10.1109/ACCESS.2016.2594194.