

Reading ACL Paper

Title: Automatic Identification and Classification of Bragging in Social Media

Authors: Mali Jin¹, Daniel Preotiuc – Pietro², A. Seza Dogruöz³, Nikolaos Aletras¹

Affiliations: University of Sheffield¹, Bloomberg², Ghent University³

The purpose of work and study performed throughout the paper was aimed to identify and classify tweets from social media and determine whether it could be consider bragging. Bragging can be seen as a form of speech that is used to make oneself or subject seem favorable. Bragging is especially common in social media where users use a platform to create an idealized self-image of themselves. This is because being viewed positively is a key driver in human behavior. The Paper was created to build on previous research involving linguistic and pragmatic to create the first large scale study regarding bragging in computational linguistics. The study involved the evaluation of different transformer models that are inject with linguistic information.

The two types of models that were evaluate consisted of binary and multi-class classifications. The multi-classification included seven classes such as achievement, action, feeling, trait, possession, affiliation, and not bragging. The purpose of including multiclassification with various bragging types was to understand how different types of bragging could either help or hurt one's self-presentation. Since bragging hasn't been studied in such a large scale regarding computational linguistics. This study was especially impactful to help linguist in better understating the context and type of bragging as well as helping social scientist analyze the relationship between bragging and personality traits, online behavior, and possible communication strategies.

Priyesh Patel
Sameer Haider

Dogruoz has been active in developing NLP techniques to help improve machine translation quality. Much of her prior work focuses on addressing the challenges of translating idiomatic language. In addition, she has also worked on the detection of sarcasm in social media posts. Aletras has been active in developing machine learning models for text classification and sentiment analysis. Much of his prior work focuses on text classification with legal and political documents. He has also worked on text generation for summarization and paraphrasing.

Both authors have worked on joint studies before as well such as their study called "Hate Speech Detection with Comment Embeddings" that worked on improving the detection of hate speech in social media. They also proposed a novel neural architecture in their paper "Sarcasm Detection on Twitter: A Behavioral and Language Model Approach" which accurately detected sarcasm in twitter posts. Other than these primary authors, Mali Jin and Daniel Preoțiuc-Pietro have also done extensive research in the NLP field.

[Describe the unique contributions of this paper](#)

With their studies and this detailed paper the authors were able to bring contributions that can help the connect previous work in pragmatic and the computational study of speech acts. One contribution they have provided to the public is a data set that contains almost seven thousand tweets that have been annotated with information that classified a tweet as bragging and if it is bragging so what type of bragging it falls into. They also performed extensive experiment with transformer models with the combination of linguistic features relating to bragging. The study results in two main classification models such as binary and multi-class. The multi-classification model is especially impactful for linguists and social scientists to study how type of bragging can fit in a spectrum of overall positive or negative bragging. Finally, the study

Priyesh Patel
Sameer Haider

and authors also provided a qualitative linguistic analysis of bragging from tweets and the model behavior in predicting bragging.

Describe how the authors evaluated their work

They test binary classification on the bragging tweets dataset which they carefully curated. They compare results between different models such as *BERTweet*, *Logistic Regression*, *BERT*, and *RoBERTa*. They also repeat this with multi-class classification among the mentioned models. The multi-class classification groups “bragging” into different categories like achievement etc. In addition to analyzing the results obtained from the models, the authors also analyze the linguistic features of the experiment. This includes unigrams, LIWC, and part-of-speech (POS) tagging associated with the bragging tweets in the dataset. In addition to this, they also analyzed how “bragging” tweets correlate with popularity of the post. Finally, they undergo error analysis to examine the limitations of their approach.

Importance of Citations

Nikolaos Aletras – 3234 (advisor)

A.S. Doğruöz - 1012

Mali Jin – 25 (PhD student)

Daniel Preotiuc-Pietro - 3014

Nikolaus Aletras has many citations to his record. This shows how impactful his work has been in the field of NLP research. He has brought upon revolutionary methods of text classification. One example of his impact is when he was able to achieve state-of-the-art performance on predicting the legal outcomes of the European Court of Human Rights cases. Aletras has been crucial to the development of NLP techniques, especially in text classification. The number of his citations serves as proof of his great work.

Priyesh Patel
Sameer Haider