Examining Word Representations between #BlackLivesMatter Movement and its Counter Protests: 2013 to 2020

Jamell Dacon and Jiliang Tang

Keywords: Digital Activism, Word Shifts, Word Representations, Sentiment Analysis

Extended Abstract

Since the fatal shooting of 17-year-old Black teenager Trayvon Martin in February 2012 by a White neighborhood watchman, George Zimmerman in Sanford, Florida, there has been a significant increase in digital activism addressing police-brutality related and racially motivated incidents in the United States. #BlackLivesMatter has become the most influential hashtag to date and is used as a “call to action” to protest police-brutality related and racially motivated related incidents and amplify social awareness surrounding Black lives [1]. In this work, we exploit social media as an authoritative tool to examine three mediums, namely the Black Lives Matter (BLM) movement and its counter protests All Lives Matter (ALM) and Blue Lives Matter (BLM). Hence, we conduct a word-level text analysis on 36,984,559 tweets to investigate users' discourse to understand the impact of digital activism within each social movement.

Motivation: Our motivation is to assess the behavioral relationships of social activists within each movement to identify similarities and discrepancies in users' discourse such as word. The main goal of our study is to understand the impact of digital activism regarding word use and determine which words best capture and contribute to each movement between February 6th, 2013, and June 30th, 2020.

Method: First, we collect currently available tweets from [2] s.t. our gathered dataset consists of 37 million tweet IDs that contain one or more of the three social movement hashtags: #BlackLivesMatter, #AllLivesMatter and #BlueLivesMatter, and use each hashtag as a proxy to split the dataset into three respective corpora. Next, we apply weight word shift graphs [3] which are a visual computational analysis tool that provides a detailed lens into textual shifts displaying fine-grained differences between corpora according to several measurements such as word frequency, entropy, sentiment, etc. Then, we implement a labMT lexicon, a general-purpose sentiment dictionary of the 5,000 most frequently used words from Twitter, New York Times, Google Books, and music lyrics constructed by the Computational Story Lab at the University of Vermont where words were rated on a continuous scale from 1 to 9, where 1 is the least happy and 9 is the most.

Results: The sentiment of a tweet is crucial to examine if the thoughts, feelings, and opinions conveyed by the activist are negative, neutral, or positive. In Figure 1, we illustrate the sentiments for the the top fifty contributing words to the difference in sentiment. As each word contribution is now the product of the difference between the word score and the reference score, and the difference between relative frequencies we see that the BlackLivesMatter vs. AllLivesMatter graph has lesser positive words compared to the BlackLivesMatter vs. BlueLivesMatter graph. This may be a result of these protest movements having a history of possessing difficult conversations to enact social change in new and complex ways. For example, words that are more negative found in these graphs are “racist”, “shot”, “war”, “suspect”, “gun”, etc. However, positive words that contribute to each corpus are “respect”, “god”, “hero”, etc., (see Figure 1).
Nevertheless, the word “protest” is characterized as negative, although being ranked in the top two contributing words in each graph, despite each movement being a protest movement. Thus, by applying a novelty visual representation, we quantify which words contribute to a difference between BLM and its counter protests. During the conference we intend on discussing more experiments such as topic modeling and thematic analysis which were also conducted to investigate complex and enormous subjective qualitative data by identifying universal conversational themes, topics, patterns and/or ideas.

References

