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

Jamell Dacon¹ and Jiliang Tang¹
¹Data Science and Engineering Lab, Michigan State University
{daconjam, tangjili}@msu.edu

Introduction

Since the fatal shooting of 17-year-old Black teenager Trayvon Martin in February 2012 by 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.

In this work, we conduct a word-level text analysis on 36,954,559 tweets to investigate users’ discourse to investigate word use to better 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 social movement.

Goals & Research Questions

The main goal of our study is to understand the impact of digital activism within each SM and determine which hashtags, words and topics capture and contribute to each movement.

1. Which words and/or tweets contribute to a difference between corpora?
2. Do these words and tweets embody the goals of each movement, or simply mitigate and suppress high-profile police-brutality related and racially-motivated incidents?

Data &Methodology

First, we collect currently available tweets from [1] 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. We then use each hashtag as a proxy to split the dataset into three respective corpora. Next, we apply weighted word shift graphs [2] and conduct thematic analysis which are visual computational analysis tools that provides a detailed lens into textual shifts displaying fine-grained differences between corpora according to several measurements such as word frequency, similarity, sentiment, etc.

Table 1. Descriptive counts for our collected dataset and each keyword, respectively. ISO 639-1 Language codes: en = English, fr = French, ja = Japanese, es = Spanish, pt = Portuguese, th = Thai, de = German, it = Italian, nl = Dutch, ko = Korean, tl = Tagalog, hi = Hindi, Haitian Creole. Note that tweets may contain one or more of the keywords and thus, being added to multiple rows.

<table>
<thead>
<tr>
<th>Tweets</th>
<th>Users</th>
<th>Retweets</th>
<th>Top Languages</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>36,954,559</td>
<td>7,610,043</td>
<td>en, fr, ja, es, pt, th, de, it, nl, ko</td>
</tr>
<tr>
<td>BlackLivesMatter</td>
<td>29,173,455</td>
<td>5,665,400</td>
<td>en, fr, ja, es, pt, th, de, it, nl, ko</td>
</tr>
<tr>
<td>AllLivesMatter</td>
<td>2,088,326</td>
<td>721,741</td>
<td>en, es, nl, ja, fr, de, pt, it, th, it, hi</td>
</tr>
<tr>
<td>BlueLivesMatter</td>
<td>2,830,532</td>
<td>418,485</td>
<td>en, es, fr, pt, nl, it, ja, de, ht, tl</td>
</tr>
</tbody>
</table>

Figure 1: Logged daily frequencies of tweets.

Figure 2: Word shift graphs using the LabMT sentiment dictionary [2] showing the top fifty contributing between two corpora with BlackLivesMatter vs AllLivesMatter on the left, and BlackLivesMatter vs BlueLivesMatter on the right.

Conflicts of Interest and Ethics

All authors declare no conflict of interest and that all data used is publicly available and distributed within Twitter’s Terms of Services. The authors state that they also have no known conflicting commercial interests or personal ties that may have impacted the work developed in this paper or may even be said to have influenced the work presented in this publication.

References
