A Twitter Opinion Mining Gold Standard for Brazilian Uprising in 2013

Tiago Cruz de França[1,2], José Orlando Gomes[2], Jonice Oliveira[2]

[1]Department of Mathematics – Federal Rural University of Rio de Janeiro (UFRRJ)
Seropédica – RJ – Brazil

[2]Postgraduate Program in Computer Science (PPGI) – Federal University of Rio de Janeiro (UFRJ)
Rio de Janeiro, Brazil

tcruzfranca@ufrrj.br, joseorlando@nce.ufrj.br, jonice@dcc.ufrj.br

Abstract. *Social media provide valuable sources of information, produced by real people about real world events, such as demonstrations or protests. The retrieval and extraction of information from such data presents some challenges, such as the production of good data sets, useful for support analysis using machine learning approaches, for example. In this paper, we present a sentiment-annotated Twitter gold standard for the analysis of Brazilian protests in 2013. The data set consists of 4,422 Twitter messages (tweets) annotated by three raters with information about the sentiment expressed in the messages. This is a valuable resource for social media-based sentiment analysis in the context of protest events.*

1. Introduction

Social media are a source of data and information about different subjects. França et al. (2014) presented some observations concerned with the data available in those media. The authors cited that the volume of data used in those media reached into petabytes. Social media provides a platform for shared information in a population, allowing them to organize themselves, to claim their rights also making it possible to raise public awareness and knowledge of such events [Hernandez and Spiro 2013].

Further to that, it provides the support and means necessary to extract information from the data and then get an understanding of events. For instance, during events like the Brazilian street protests, people were motivated to share their opinions [França and Oliveira]. In these settings, extracting the opinions expressed in such messages is fundamental to providing insights, enabling an analysis of the underlying political processes and dynamics [Hürlimann et al. 2016]. Machine learning approaches make it possible to analyze quantities of data which would be impossible for humans to analyze.

Automatically analyzing messages written in languages such as Brazilian Portuguese, can be harder than in other languages. For example, in the context of natural language process and sentiment analysis there are no adequate tools or a sentiment lexicon for Brazilian Portuguese, such as there are for English [Freitas 2015, Araújo et al 2016]. So, in such scenarios, where messages are not written in English, the use of machine learning approaches are often a good choice [França and Oliveira 2014]. In order to enable the sentiment analysis (or opinion mining) using machine learning it is necessary to provide annotated and reliable data sets [Saif et al. 2013, Hürlimann et al. 2016]. The annotation of those data sets are difficult procedures and there aren't many data sets available, there are even less data sets of messages in Brazilian Portuguese [França and Oliveira 2014].

This work presents a gold standard human annotated data set. The data set consists mainly of tweets written in Brazilian Portuguese. A random sample was built upon 432,975 messages related to Brazilian protests in 2013. The data set has 4,422 tweets. Three annotators assigned just one class to each tweet in the data set. The possible classes were positive, negative or neutral sentiment about the protests. We also present some agreements and inter-rater agreement measures within the data set. The data set is available in https://labcores.github.io/p_tcruzfranca/.

The remainder of this work is organized as follows. In section two, we have presented some related works. In section three, we have described some information about the events to which the data set is related. Section four has the description of the data sets, the method and consolidation of the gold standard data set. Finally, section five has the final considerations.

2. Related Works

There are some existing tweet data sets available [Hassan et al. 2013]. For instance, those providing information to analyze data about the Brexit. Two examples are the Twitter gold standard for Brexit referendum [Hürlimann et al. 2016] and the #ImagineEurope project [ImagineEurope 2013a; ImagineEurope 2013b]. They both collected a data set of tweets using hashtags as filters. In general, those data sets were collected in the time leading up to the referendum (at least in the same year, months or weeks before the referendum day).

Hürlimann et al. (2016) selected certain hashtags and user citations to get data related to the Brexit. They then took a random sample of the English message data to be annotated. The main annotation task that they performed annotated two thousands tweets into five classes. Subsequently, they calculated certain agreements and inter-rater agreements in relation to the data. This proposal intends to work in similar way, but we have annotated a predominantly Brazilian Portuguese data set, related to protests in Brazil.

There are also other data sets that exist related to the Brexit or other real events (such as political debate, discourse or discussion, for example). In Ontotext (2016) 1.5 million tweets were utilized, while [Sheth 2016] checked messages posted to support the staying in or not (the leaving) of the Europe Union, three hours before and four hours after the referendum. Priego [Priego 2016a] and Priego[Priego 2016b] also studied tweets related to that event, in order to analyze the messages supporting, the staying in or leaving of, the Europe Union.

Go et al. (2009) introduced the Stanford Twitter Sentiment Corpus which consists of two different sets; a training and a test. On the one hand, the training data set has 1.6 million of tweets labeled as either positive or negative. The labels were assigned in a automated way using messages with emoticons (e.g. =(or =D). Despite the amount of messages, there is no guarantee that the emoticons really represent the sentiment of the message. On the other hand, the test data set is manually annotated and contains 177 negative, 182 positive, and 139 neutrals tweets.

Another data set was constructed to analyze the Obama-McCain debate during USA elections in September 2008 [Shamma, Kennedy and Churchill 2009]. This data set has 3,238 tweets annotated using Amazon Mechanical Turk, utilizing at least three annotators for each tweet. The classes assigned were positive, negative, mixed or other.

Another example is the Sentiment Strength Twitter Dataset which consists of 4,242 tweets manually labeled with positive and negative sentiment strength [Thelwal, Buckey and Paltoglou 2012]. Rieser (2014) presented a gold standard annotated data set of Arabic tweets which contains 8,868 messages.

Hernandez and Spiro (2013) also investigated tweets related to the Brazilian protests in 2013. However, the authors did not share their data set. Theodoro et al. (2015) utilized Twitter data to identify influencing factors in that social media. França and Oliveira (2014) analyzed the sentiment expressed in tweeted messages related to the Brazilian protests. This related work was a study using the preliminary annotated data of our data set. Such a preliminary data set has not been published before. We do not know of any other data set of Brazilian Portuguese related the Brazilian protests in 2013.

3. Data Set Contextualization

In 2013 Brazilians were facing huge protests and social media, such as Twitter, were used by the public as a channel to express opinion about the demonstrations. These people could have been the activists directly involved in protests or everyday citizens just expressing an opinion on events without actually participating in them. During this period certain expressions (words and hashtags) associated with the protests became very common (entering the list of trending topics in social media generally). This meant that some users utilized words employed in the protest context, although their intentions were not directly related to events (i.e. advertisers/self publicists).

We will describe the method used to collect the data and to build the data set, in relation to the protests, in section 4. However, to understand the data set, in this section we have made a description of the main events related to the protests. Brazil is a huge country with practically continental dimensions and is one of the 10 largest economies in the world, moreover, it has a population of more than 200 million people, spread about in different national regions. Moreover, more than 102 millions of it's inhabitants access the Internet, among which 89% use a smartphone to access social media [Portal Brasil, 2016], in order to to share their opinion about different subjects [Hernandez & Spiro, 2013].

The protests that occurred in Brazil in 2013 were compared to the protests that happened during the impeachment of the former President Color in 1992. But in 2013, Brazilians complained about increases in public transport ticket prices, as well as government corruption, amongst other things. Indeed, since 2012, Brazilians had marched against increases in public transport fares. Despite this, only in Jul 2013 did the major media channels report events, amidst claims of violence involving both the protesters and the police that clashed with them [Datafolha 2013c]. During these demonstrations, a number of violent events took place. For instance; damage to public and private buildings, the throwing of Molotov cocktails and many people were injured by rubber bullets, tear gas and incendiary incidents. At least six deaths were associated with the protests. The last major violent clash between security forces and the population had occurred in 1968 during protests calling for the end of military dictatorship in Brazil [Brazil, 2013].

Figure 1 summarizes events that occurred between June 1 and July 31, 2013. We have focused on that interval because during these months the protests were at their most intense. The summarization of events is presented as a timeline and the information is based on the following references [Brazil 2013]. Initially, the marchers complained about rises in public transport fares. Motivated by the large numbers of protesters and the police clashes, the demonstrations immediately gained attention in social media and major Brazilian media outlets published some information on the events happening. It is important to highlight that from the 15th until 30th of June, Brazil hosted the FIFA Confederations Cup and the World Youth Days from 23th to the 28th of July.

After a series of demonstrations in Brazil the nature and number of the public demands needed to be clarified. Besides ticket transport prices, the people also demonstrated their disappointment regarding; government corruption, certain legislative proposals (e.g. PEC 37) and the large expenses incurred by events that were to be held in the country. There was no official social leadership, with neither political party leading the protests in Brazil during this time. Therefore, a group named Anonymous Brazil published a video promoting 5 main issues. We have highlighted two of them. Firstly, a rejection of a proposed amendment to the Brazilian constitution known as PEC 37 which intended to prohibit investigations by the federal public ministry. The PEC 37 was understood to be a way of facilitating politician's departures, without them paying for their corruption. Secondly, they called for the opening of an investigation for verifying irregularities related to money spent on infrastructure for the FIFA World Cup. An event that would take place in 2014.

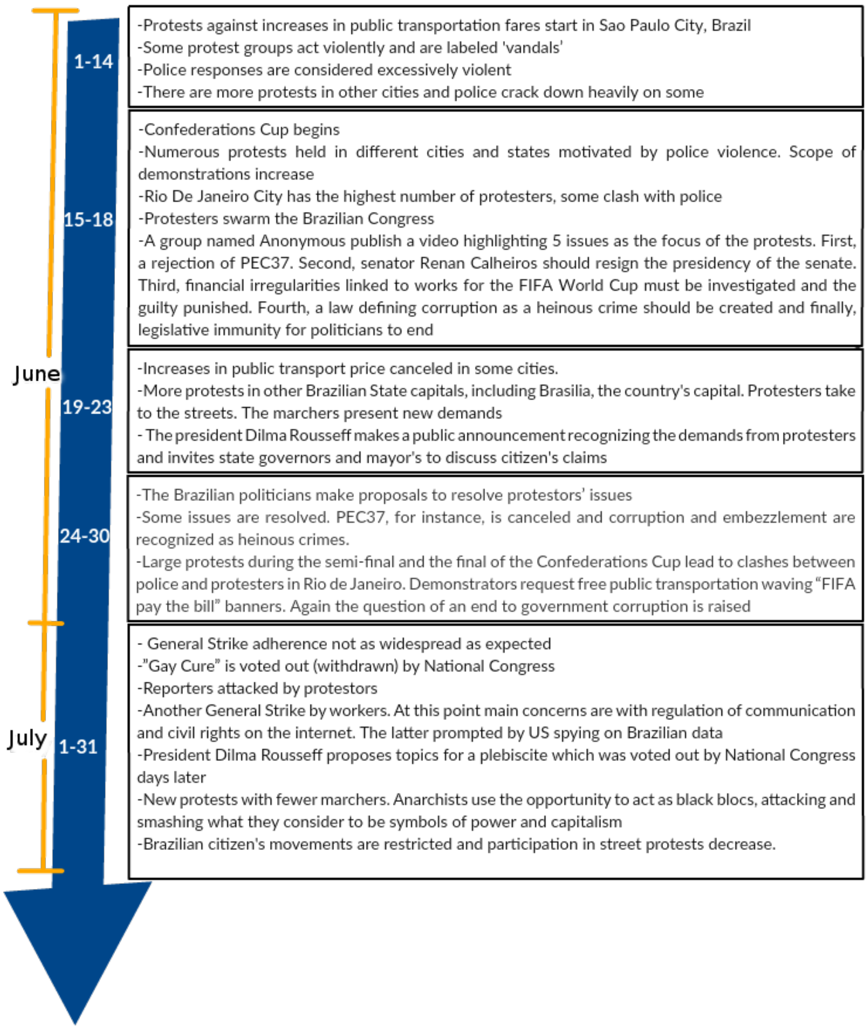

Figure 1. The Timeline of the main events during the Brazilian protests, 2013

Certain Brazilian agencies conducted surveys related to the protests. An opinion poll showed that 55% of citizens from the city of São Paulo had agreed with the protests on June 14 [Datafolha, 2013d]. According the same survey, 41% of respondents had disagreed, while 4% had no opinion or were indifferent. In other words, 55% of citizens had a positive opinion about protests while 41% had a negative opinion. A week later, on June 21st, 66% still agreed with the protests, even though just 2 days earlier public transportation fares had been reduced [Datafolha, 2013e]. On October 28th in São Paulo city data showed that the support for the protests had decreased from 89% to 66% in October while the rate of disagreement grew from 8% to 31% [Datafolha, 2013a]. In February 2014 the interviewees considering the protests to be positive reduced to 52% while those who considered them to be negative increased to 42% [Datafolha, 2014].

A nationwide public opinion survey was made on June 24th, 2013 by the Brazilian Institute of Public Opinion and Statistics Institute (IBOPE). The survey results had shown 75% Brazilians agreed with the demonstrations and 22% had disagreed [IBOPE, 2013a]. IBOPE also asked Brazilians about political party representation and 89% of the people answered that no political party represented them [IBOPE, 2013b].

Other surveys demonstrated the importance of social media in the Brazilian protests. The results of [Datafolha, 2013b] showed that on June 19th, 93% of the activists organizing the demonstration through social media considered such media as their main source of information. According to IBOPE [IBOPE, 2013b] results on June 20th, 62% of marchers knew of the protests through Facebook and 75% of people invited other people using Facebook and Twitter.

4. Method and Data Set Consolidation

Our aim is provide a random sample to support the opinion mining (sentiment analysis) from tweets related to Brazilian protests in 2013. Many demonstrations occurred during this year in Brazil between June and August, which represents a time period relating to events as summarized in section 3.

The first step in data retrieval was to define some filters. During the protests we observed the content posted relating to the events in which we are interested. Then we defined some hashtags used to index the content published, as expressions associated with the protests. Secondly, we kept in mind that we wanted to understand the opinion of the users from Twitter. So we defined a protocol to assist in our analysis. The protocol comprises of a defining guide question and some possible answers (the classes) since we are interested in support supervised machine learning analysis to extract information from the data set.

In general, there are two basic ways to get posts from Twitter, one is using the Twitter API and the other is using Twitter stream[1]. Both approaches are provided by a REST API and require some credentials from a user to access resources. We have used the Twitter API to build our raw data set. That API imposes some restrictions, such as limits of requests and limits of time (posts newer than 7 days) to query and deliver data.

4.1 Sample and Filtering (Data Collections)

In regards to data collection, we used the Twitter's API during the Brazilian protests between June 1 and July 31, 2013 (inclusive). In order to cover posts related to the protests, we utilized 22 hashtags used by the activists which are presented in List 1. The criteria chosen for those hashtags was based on the manual identification of common keywords associated with the protests events in Brazil during that period.

1 https://dev.twitter.com/overview/api

Each tweet is returned by the API as a JSON representation. That JSON comprises all metadata provided by Twitter. Those metadata include, for example, data about the user that published the message, a timestamp, possibly a geolocation and so on. Those tweets formed a raw data set with 432,975 messages published by 180,005 different users (approximately 2.41 tweets per user). The information contained in the raw data set is summarized in Table 1. Figure 2 depicts the frequency of tweets per day between the Jun 1 and Jul 31, 2013. We should point out that most messages were written in the Brazilian Portuguese language.

#acordabrasil, #vemprarua, #ForaFifa, #ogiganteacordou, #anonymousbrazil, #MPL, #passelivre, #pec37, #mudabrasil, #ChangeBrazil, #anonymousbrazil, #protesto, #foradilma, #protestorj, #protestabrasil, #primaverabrasileira, #forafeliciano, #ocupa, #copapraquem, #protest, #pec33, #pec99

List 1. Hashtags used to retrieve Twitter's messages

Table 1. Basic raw data set information

	# Amount
Total of messages	432,975
Total of users	180,005 (average of ~2.41 messages per user)
Period of data collect	61 days (from Jun 1 up to Jul 31, 2013)

We build a random sample of 1% of messages for each day from the stream represented by the raw data set. In other words, for each day we randomly chose 1% (with a minimum of 1 message in a day) of data to be labeled. The total amount of messages was 4,422 tweets.

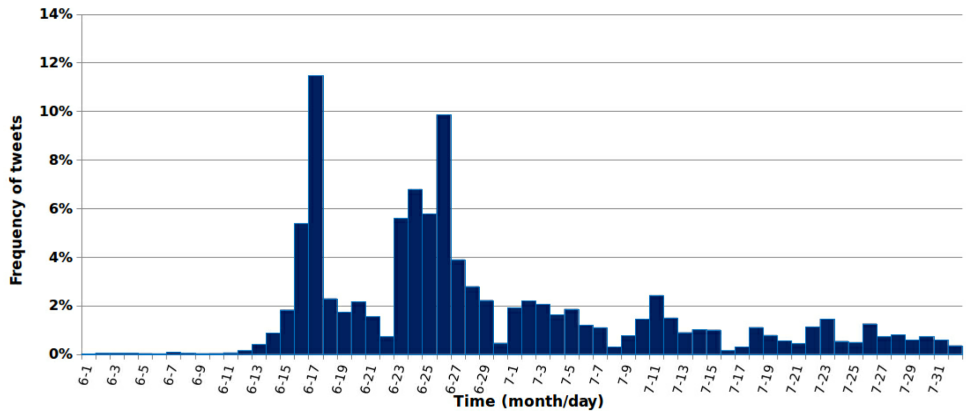

Figure 2. Frequency of Tweets per Day

4.2 Annotation

The 4,422 tweets sampled were presented to three volunteer annotators. Before starting the annotation task, each annotator received basic explanations related to both "why" and "when" the dataset was built. Then, the following two explanations were presented. The first one was concerned with the motivation for understanding how the activists and population expressed their opinion (or sentiment) in the social media (mainly on Twitter). The second, was related to the period of data retrieval that occurred during the many Brazilian protests events in 2013.

The aim was to classify whether the opinion expressed in the messages was in some way positive or negative, regarding the protests. To do that, we defined a **key question** as a guide for the classification task. The question states *"(does) the tweet text expresses an opinion of agreement (positive), disagreement (negative) or neither (neutral) to the protests?"*. Only one answer should be assigned to each message. Thus, the annotators judged whether a message expressed an opinion in favor (positive) or against (negative) the protests. If a classifier judged that a message didn't express a positive or negative opinion for any reason, that message would be considered neutral. In other words, **we have defined three (answer) classes**: **positive, negative or neutral.**

We should note that the annotation tasks were started eight months after the protests events. At first, a pilot was built with only 200 messages [França and Oliveira, 2014]. After that, the sample was built in a random way taking 1% of messages for each day of the raw data set. We also assume a minimum of one message per day, if the 1% represents a value less than one. The entire labeled data set was then finished, one and half years after the protests. The results of the annotation task are in Table 2. This approach enables us to get some important feedback, such as that the majority of the messages were in Brazilian Portuguese. Furthermore, we could also take some inter-rater agreement measures between the annotators.

Table 2. Rating among annotators and classes

Classes	Annotators-1	Annotators-2	Annotators-3
Positive	1,786	1,723	1,945
Negative	191	277	294
Neutral	2,445	2,422	2,183
# Total of Labeled Messages		4,422	

4.3 Agreement

The approach of classifying the messages three times with different annotators enables us to verify the inter-rater agreement among such raters. The inter-rater agreement aggregates information about how reliable the annotated data set is. That means low inter-rater agreement could produce unreliable results. This could happen because of problems in annotation method, during the annotation task, or in the data set.

Table 3 shows the distribution of messages with regard to the number of annotators opinion rater agreement, providing a different view of agreement. In Table 4 we present two inter-rater agreement metrics (Fleiss' Kappa and Krippendorff's alpha) for the sentiment annotations. Fleiss' Kappa is useful when all messages are labeled by the annotators. Krippendorff's alpha allows the inter-rater agreement to be calculated even if not all the data is labeled for all volunteers.

Table 3. Agreement among the annotators (raters)

	# tweets	%
Unanimous sentiment	3,449	~ 78%
Two sentiments	936	~ 21%
Three sentiments	37	~ 1 %
Total	4,422	100%

Table 4. Inter-rater agreement among raters

Statistic inter-rater agreement	Result
Fleiss' Kappa	73%
Krippendorff's alpha	65%

If inter-rater agreement is low or negative, then someone could conclude that an automated supervised classification task is not useful because it's result is not reliable. That is, the results of an automated classification will reflect the manually annotated task. Thus, it will not represent non-valuable information to understand real case scenarios. In our case, we argue that inter-rater agreement values enable someone to know how reliable the classification is and thus how valuable the data set is.

There are some points about both Kappa value and Krippendorff's alpha that we want to highlight. Landis (1977) had proposed a scale in which Fleiss' Kappa greater than 0.6 would be good values, between 0.41 and 0.6 moderate, and less than 0.41 fair, slight or poor when less than 0. However, there is no widely agreed acceptance about the interpretation of Kappa results [Gwet 2014]. Besides that, when the number of classes are fewer, the kappa tends to be higher [Sim and Wright 2005].

Krippendorff's alpha presents some disadvantages when used to measure nominal data. It is more general than Fleiss' Kappa, but has some weaknesses such as it ignores how many times a subject was annotated. Another consideration is that just the disagreement is evaluated while the agreement is ignored. In other words, just the disagreements define the result. Finally, the disagreements are larger for nominal data leading alpha to a small result.

4.4 Data Set Consolidation and Description

The gold standard was obtained after a consolidation on the annotations presented in Table 2. That consolidation procedure follows two steps. We use 1) a majority vote for the sentiment, and 2) when the raters disagreed entirely (each one assigned three different sentiments to a tweet), then the decision was assign the tweet as neutral. As presented in Table 3, a small amount of messages were assigned in three different class (just 37, less than 1%).

Table 5. The gold standard final consolidation

	# consolidation	%
Positive	1752	~40%
Negative	198	~4%
Neutral	2472	~56%
Total of Labeled Messages (#)	4422	

Unfortunately, according to Twitter policy[2] a third party cannot provide entire tweet objects, only the tweet's IDs which in turn must be shared as a spreadsheet or PDF file. Therefore, the raw data set is available for download as a list of IDs. We also provide a python script to retrieve tweets by IDs using the Twitter's API. The resource API utilized enables the recovery of posts older than 7 days. Therefore, although the data set being composed by messages are unrecoverable using Twitter API search query, it is possible when the tweet's IDs are known.

The shared data set has the tweet IDs, the assigned classes by humans in three distinct column and the consolidation column. The classes were assigned as follows; messages classified as *positive* has received an uppercase "P" as mark. Those classified as *negative* or *neutral* received an uppercase mark as either "N" or "NN", respectively. Similarly, the raw data set is available for download as a spreadsheet, but containing just the tweet IDs.

Most of the messages were classified as being neutral. It should be noted that neutral messages are more common among social media messages than in the surveys presented in section 3. This is because the neutral class groups together all messages that did not present positive or negative ratings. This means that any result which does not represent feeling, will be classified as neutral. For example, a message reporting an event or an advertising message indexed with one of the hashtags adopted in the project could be used to gain attention, for example. In addition, IBOPE and Datafolha question users, which is different from extracting their opinion from a text that is not the answer

2 https://dev.twitter.com/overview/terms/policy

to a pre-prepared question, but rather a free expression by the user. In the Gold Standard data set, the positive and negative opinions in June were 34% and 5% respectively. In July, the positive was 50% and the negative 4%.

Table 6. Examples of tweets and annotations assigned

	Tweet text	Assigned sentiment
1	Falta de leitos nos hospitais públicos leva milhares de pacientes recorrem ao Ministério Público para se tratarem #VempraRua #SOSSUS	Positive
2	Cara, desisto de fazer vocês entenderem. Não adianta ficar nessa mobilização " #mudabrasil" sem propôr mudanças reais.	Negative
3	O pau comendo em várias cidades do Brasil e o Fantástico começa com música feliz, falando sobre a Copa das Confederações. #protestorj	Neutral
4	Convoco lutadores de Jiu-Jitsu e Muay Thai a participar dos protestos para dar surra nos baderneiros e saqueadores de lojas. #mudabrasil	Negative (it received two negative and one positive assignment)
5	A torcida continuou cantando o hino nacional, mas n foi p fazer bonito para a rede Globo e sim p mostrar sua forca! #OGiganteAcordou	Neutral (received the three possible sentiments annotations)

5. Final Considerations

In this work we have described a Gold Standard data set for the Brazilian protests in 2013. Besides the data set description, we also presented the method, consolidation and some measures of agreement and inter-rater agreement among three human annotators. The annotated data set provides a resource for observing the social and discourse dynamics associated to the protests in Brazil between June 1st and July 31st, 2013.

The messages present in the data set were classified in one of three classes: positive, negative or neutral. Those tags represent whether the messages present a agreement or disagreement about the protests. If no sentiment could be assigned for any reason, then the message was classified as neutral. After the human annotation task, we defined two basic consolidation decisions. The final result would be the sentiment with more assignment, or if a total disagreement occurred (each annotator assigned a different tag), the message was defined as neutral.

This data set is useful for anyone interested in exploring data about the protests, or to evaluate machine learning analysis specifically using messages written in the Brazilian Portuguese language, since our messages are mostly Brazilian Portuguese messages. The data set also supports the use or evaluation of, tools for natural process languages for Portuguese or tweet messages.

Besides the gold standard data set, we have also presented some overviews about the raw data set retrieved during the events. We named that data set the ‘raw data set’. We should point that the raw data set is a convenient sample returned by Twitter, since we have no way to define a sampling method. We knew previously that Twitter has restrictions or limits on the data retrieval and so we will use a sample without any statistical definition. Another point to be taken into consideration is the fact that many citizens have no access to or an account on Twitter. For instance, poor people often have no means to access the Internet. Therefore, even before the Twitter restrictions, social media is, in general selective.

References

Brazil (13 jul 2017). 2013 protests in Brazil. *Wikipedia*. https://en.wikipedia.org/w/index.php?title=2013_protests_in_Brazil&oldid=790394354.

Datafolha (2013 6AD). Paulistanos defendem continuidade de protestos e foco em saúde e educação - Opinião Pública. http://datafolha.folha.uol.com.br/opiniaopublica/2013/06/1300362-paulistanos-defendem-continuidade-de-protestos-e-foco-em-saude-e-educacao.shtml, [accessed on Jul 30].

DataFolha (14 jun 2013). Paulistanos aprovam protestos, mas rejeitam vandalismo e tarifa zero - Opinião Pública. http://datafolha.folha.uol.com.br/opiniaopublica/2013/06/1295431-paulistanos-aprovam-protestos-mas-rejeitam-vandalismo-e-tarifa-zero.shtml, [accessed on Jul 30].

Datafolha (14 jun 2013a). Paulistanos aprovam protestos, mas rejeitam vandalismo e tarifa zero - Opinião Pública. http://datafolha.folha.uol.com.br/opiniaopublica/2013/06/1295431-paulistanos-aprovam-protestos-mas-rejeitam-vandalismo-e-tarifa-zero.shtml, [accessed on Jul 30].

Datafolha (28 oct 2013b). Apoio às manifestações cai de 74% para 66% - Opinião Pública. http://datafolha.folha.uol.com.br/opiniaopublica/2013/10/1363246-apoio-as-manifestacoes-cai-de-74-para-66.shtml, [accessed on Jul 30].

FRANCA, T. C. and Oliveira, J. Proceedings of III Brazilian Workshop on Social Network Analysis and Mining (BraSNAM), Brasilia. Anais do Congresso da Sociedade Brasileira de Computação, 2014.

França, T., Faria, F., Rangel, F., Farias, C. and Oliveira, J. (2014). *Big Social Data: Princípios sobre Coleta, Tratamento e Análise de Dados Sociais*. Curitiba-PR: Bernadette Farias Lóscio.

Go, A., Bhayani, R. and Huang, L. (2009). Twitter Sentiment Classification using Distant Supervision. *Processing*, p. 1–6.

Hürlimann, M., Davis, B., Cortis, K., et al. (2016). A Twitter Sentiment Gold Standard for the Brexit Referendum. In *Proceedings of the 12th International Conference on Semantic Systems*. , SEMANTiCS 2016. ACM. http://doi.acm.org/10.1145/2993318.2993350.

IBOPE (24 jun 2013a). 75% dos brasileiros são favoráveis às manifestações públicas. http://www.ibopeinteligencia.com/noticias-e-pesquisas/75-dos-brasileiros-sao-favoraveis-as-manifestacoes-publicas/, [accessed on Jul 30].

IBOPE (25 jun 2013b). 89% dos manifestantes não se sentem representados por partidos. http://www.ibopeinteligencia.com/noticias-e-pesquisas/89-dos-manifestantes-nao-se-sentem-representados-por-partidos/, [accessed on Jul 30].

ImagineEurope (2017). Building a Twitter Dataset to Find out How People View the EU (with images, tweets) ·. https://storify.com/ImagineEurope/building-a-twitter-dataset-to-find-out-views-on-th, [accessed on Jul 30].

ImagineEurope ([S.d.]). EU Twitter Sentiment Analysis (with images, tweets) · ImagineEurope. https://storify.com/ImagineEurope/initial-sentiment-analysis, [accessed on Jul 30].

Monroy-Hernandez, A. and Spiro, E. (2013). Research Notes » Blog Archive » How is the Brazilian Uprising Using Twitter? . http://blogs.harvard.edu/andresmh/2013/07/how-is-the-brazilian-uprising-using-twitter/, [accessed on Feb 4].

Ontotex (2016). White Paper: #BRexit Twitter Analysis. *Ontotext*. https://ontotext.com/white-paper-brexit-twitter-analysis/, [accessed on Jul 30].

Portal Brasil (14 jun 2013). Maioria da população é a favor dos protestos, mostra Datafolha - 14/06/2013 - Cotidiano. http://www1.folha.uol.com.br/cotidiano/2013/06/1294919-maioria-da-populacao-e-a-favor-dos-protestos-mostra-datafolha.shtml, [accessed on Jul 30].

Portal Brasil (2016). Pesquisa revela que mais de 100 milhões de brasileiros acessam a internet — Portal Brasil. http://www.brasil.gov.br/ciencia-e-tecnologia/2016/09/pesquisa-revela-que-mais-de-100-milhoes-de-brasileiros-acessam-a-internet, [accessed on Jul 30].

Priego, E. (21 jun 2016a). “Stronger In”: Looking Into a Sample Archive of 1,005 StrongerIn Tweets. *Ernesto Priego*. https://epriego.wordpress.com/2016/06/21/stronger-in-looking-into-a-sample-archive-of-1005-strongerin-tweets/, [accessed on Jul 30].

Priego, E. (21 jun 2016b). “Vote Leave”: Looking Into a Sample Archive of 1,100 vote_leave Tweets. *Ernesto Priego*. https://epriego.wordpress.com/2016/06/21/vote-leave-looking-into-a-sample-archive-of-1100-vote_leave-tweets/, [accessed on Jul 30].

Refaee, E. and Rieser, V. (2014). An Arabic Twitter Corpus for Subjectivity and Sentiment Analysis. . European Language Resources Association. https://researchportal.hw.ac.uk/en/publications/an-arabic-twitter-corpus-for-subjectivity-and-sentiment-analysis, [accessed on Jul 30].

Saif, H., Fernández, M., He, Y. and Alani, H. (2013). Evaluation datasets for Twitter sentiment analysis: a survey and a new dataset, the STS-Gold. . http://www.di.unito.it/~patti/essem13/index.html, [accessed on Jul 30].

Shamma, D. A., Kennedy, L. and Churchill, E. F. (2009). Tweet the debates: Understanding community annotation of uncollected sources. In *In WSM ’09: Proceedings of the international workshop on Workshop on Social*.

Sheth, A. (24 jun 2016). #Brexit: “there is a big trouble for #remain” — Some Lessons from Real-time #socialmedia Analysis. https://www.linkedin.com/pulse/brexit-big-trouble-remain-some-lessons-from-real-time-amit-sheth, [accessed on Jul 30].

Thelwall, M., Buckley, K. and Paltoglou, G. (jan 2012). Sentiment Strength Detection for the Social Web. *J. Am. Soc. Inf. Sci. Technol.*, v. 63, n. 1, p. 163–173.

Theodoro, I., Stearns, B., França, T. C. and Oliveira, J. (2015). Padrões de Comportamento de Usuários nos Protestos Brasileiros via Twitter. In *XII Simpósio Brasileiro de Sistemas Colaborativos (SBSC), 2015*.