#BREXIT VS. #STOPBREXIT: WHAT IS TRENDIER? AN NLP ANALYSIS

Marco A. Palomino[1] and Adithya Murali[2]

[1] School of Computing, Electronics and Mathematics, University of Plymouth, Drake Circus, Plymouth, PL4 8AA, United Kingdom

[2] School of Computing Science and Engineering, Vellore Institute of Technology, Vellore - 632 014, Tamil Nadu, India

ABSTRACT

Online trends have established themselves as a new method of information propagation that is reshaping journalism in the digital age. We argue that sentiment analysis—the classification of human emotion expressed in text—can enhance existing algorithms for trend discovery. By highlighting topics that are polarised, sentiment analysis can offer insight into the influence of users who are involved in a trend, and how other users adopt such a trend. As a case study, we have investigated a highly topical subject: Brexit, the withdrawal of the United Kingdom from the European Union. We retrieved an experimental corpus of publicly available tweets referring to Brexit and used them to test a proposed algorithm to identify trends. We validate the efficiency of the algorithm and gauge the sentiment expressed on the captured trends to confirm that highly polarised data ensures the emergence of trends.

KEYWORDS

Twitter; sentiment analysis; world clouds; text mining; information retrieval.

1. INTRODUCTION

Twitter is a microblogging service founded in 2006 that enables people to post short messages—namely, *tweets*—expressing their interests and attitudes [1]. Twitter users can communicate with each other, with groups and with the public at large; thus, tweets are often experienced by broader audiences than just the interlocutors. Commonly, Twitter users employ *hashtags*—words or phrases preceded by a hash sign '#'—to categorise tweets topically, so that others can search and follow conversations identified by a particular hashtag. A more detailed description of Twitter and its jargon can be found in [2].

As research involving Twitter continues to grow, it has become clear that tweets contain plenty of valuable information [3], ranging from tracking the effectiveness of marketing campaigns [4] to forecasting stock market variables [5]. However, the challenge of retrieving, storing and processing the colossal number of tweets available keeps mounting—the number of tweets published grew from 5,000 per day in 2007 to 500,000,000 per day in 2013—the annual volume of tweets keeps growing at approximately 30% per year [6].

Any of the 1.3 billion registered Twitter users [7] is able to create news stories. Twitter has dwarfed the ability of journalists and commentators to digest and analyse information within fixed news cycles [8]. Twitter is now known as "the place where news break first" [9]. Still, it is

Natarajan Meghanathan et al. (Eds) : NLP, ARIA, JSE, DMS, ITCS - 2019
pp. 19-34, 2019. © CS & IT-CSCP 2019 DOI: 10.5121/csit.2019.91203

impossible for a person to monitor all the tweets around the world. Thus, Twitter provides a list of *trending topics* to index and summarise on-going discussions [10].

While Twitter determines trends by a proprietary algorithm which is tailored for each individual user—based on who the user follows, what interests the user has and where the user is located [10], we want to make sense of large corpora of tweets regardless of individual users. Our objective is to automatically classify tweets into clusters to facilitate their exploration. As an initial step, we have chosen to monitor a highly topical subject at the time of writing: *Brexit*, a portmanteau of *British* and *exit*, which refers to the possible withdrawal of the United Kingdom from the European Union [11].

We will look into the use of tools to gauge the sentiment expressed in trending topics. By combining sentiment analysis with trend discovery, we expect to recognise not only what is trending, but also how the sentiment expressed will affect the rising or declining of emerging trends. In the long run, we wish to ascertain whether trending topics are characterised by a strong sentiment polarity—it appears likely that highly polarised tweets can spark off the discussion and, in turn, create a new trend.

The remainder of this paper is organised as follows: We begin with an overview of related work. Then, we describe the dataset that we used for our experimentation. Afterwards, we describe the trending topic algorithm that we implemented, and we use it to evaluate the impact of sentiment on trending topics. Finally, we present our results and conclusions.

2. RELATED WORK

Trending topics is a term coined by Twitter to refer to the most used keywords on the social network at a particular time [10]—these are keywords and hashtags that experience a surge in popularity—"what everyone is talking about". Commonly, trending topics evolve around cultural occurrences, such as current events, celebrity announcements and breaking news. Whilst there is no limit to how long a topic may stay popular, trending topics normally have a shelf-life of one day to one week [12].

Twitter determines trending topics by means of an algorithm which considers how many tweets are published on a particular topic and how much time it takes to reach such a number of tweets [10]. The algorithm favours sudden increases in the number of tweets, rather than a gradual sustained growth. Broadly speaking, a one-day growth would create a trend, while 30 days would not. Such a distinction, and the consequences of it, were explored in *The Washington Post* article, "*Why Didn't #FreddieGray Trend on Twitter?*" [13]. As the number of tweets using the hashtag `#FreddieGray` built up over time, the volume increased at the same rate of the traffic. Since there was never a "spike'' in the use of the hashtag, the topic did not trend—instead, the conversation carried on gradually during various days.

While Twitter's algorithm is proprietary and thus not open to scrutiny, detecting trending topics has been investigated by several researchers [14-17]. Petrovic et al. [17] developed an event detection approach capable of scaling over big data streams using *paraphrases*, alternative ways of expressing the same meaning—for example, the phrase "he got married" can be paraphrased as "he tied the knot". Petrovic et al. were able to detect that some tweets previously thought to be about new events were actually paraphrases of other tweets published formerly. Petrovic et al. evaluated their algorithm by getting two human annotators to label the output as being about an event or not. While we recognise the importance of this work, we are more interested in

measuring how many trends we can actually detect—and determining the sentiment expressed in such topics.

Shamma et al. [18] focused on *peaky* topics—topics that show highly localised, temporary interest. Shamma et al.'s method concentrates on obtaining "peak" terms for a given timeslot, as opposed to the whole corpus. By using the whole corpus of tweets, Shamma et al. favoured *batch*-mode processing—i.e., getting the data initially and processing it later on—which is unsuitable for the real-time analysis that we want to pursue.

Closer to our goals is the work by Benhardus [15], who uses standard statistical techniques, such as *TF-IDF* [19]. Whilst Benhardus prioritises the discovery of general trending topics—typically finding information about celebrities and new hashtags—we focus on news headlines referring to Brexit, partly because they were easier to obtain at the time of writing, but also because they were ideal to test our work, as we will explain below.

2.1.Sentiment Analysis

Sentiment analysis is the computational study of opinions and subjectivity in text [20]. The focus of sentiment analysis is the detection of *sentiment polarity*, by which the opinion of a piece of text is identified as *positive* or *negative* [21].

The main approaches employed by sentiment analysis are machine learning approach, lexicon-based approach and hybrid approach [22]. Whereas the machine learning approach uses linguistic features, the lexicon-based approach relies on a sentiment lexicon—a collection of pre-compiled terms that express sentiments. The hybrid approach combines both the machine learning and lexicon-based approach.

The number of tools developed to perform sentiment analysis is constantly increasing [23]. For the purpose of our work, we have selected two specific tools: *SentiStrength* [24] and *VADER* [25]. The reason why we have chosen two separate tools is the lack of consensus among them [26]. Hence, rather than relying on a single tool, we prefer to consider a couple of options and compare and contrast their differences.

3. EXPERIMENTAL CORPUS

The total number of tweets contained in the corpus that we used for our experiments is 62,397. Such a collection was gathered after retrieving tweets for 24 continuous hours, starting on Friday 8th March 2019 at 15:28:00 (GMT)—hereafter, all times are GMT times. The first tweet was captured at 15:28:30 and the last one on Saturday 9th March 2019 at 15:19:15.

Our experimental corpus consists of publicly available tweets labelled either with the hashtag `#Brexit` or `#StopBrexit`. The tweets were gathered in Plymouth in the UK using *Tweepy* [27], an open-source, Python library for downloading tweets in real time [28]. Tweepy makes it easier to use the *Twitter Streaming API* [29] by handling authentication and connection—a *stream listener* captured the tweets labelled with the hashtags `#Brexit` and `#StopBrexit` as soon as they were posted [27].

Figure 1 displays the volume of tweets captured per hour during our experiment—we captured 2,600 tweets per hour, on average. As shown in Figure 1, the volume of tweets started to decrease at 20:28 in the evening of 8th March 2019, and it did not increase again until after 04:28, approximately, in the early morning of the following day.

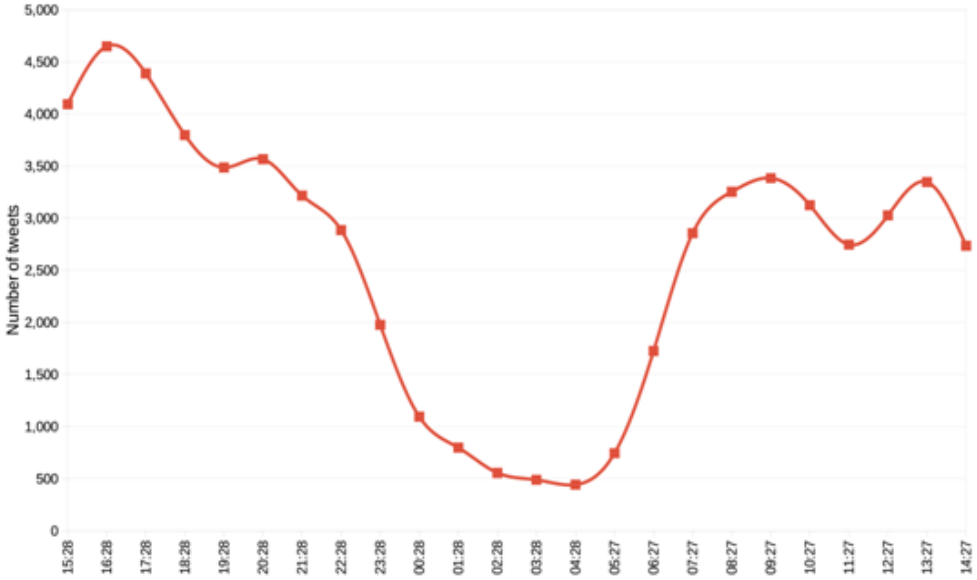

Figure 1. Volume of tweets per hour

Marketing experts typically recommend that tweets should be posted between noon and 15:00 throughout the week to improve engagement with followers [30]. Hence, we expected to find a decrease in the number of tweets published throughout the middle part of our retrieval, as confirmed by Figure 1.

We chose 8th March 2019 as the date to collect our experimental corpus, because Theresa May—Prime Minister of the UK and Leader of the Conservative Party since 2016—warned on that day that if MPs rejected her Brexit deal a week later, the UK may never leave the EU at all [31]. Theresa May told those preparing to take a decision in the House of Commons that they should move "past the bitterness'' of the debate. She also demanded the EU to give more ground in deadlocked talks--her exact words were "let's get it done" [31]. Under the then current UK political climate, we expected these news to generate a fair amount of discussion involving the hashtags `#Brexit` and `#StopBrexit`, which we could capture on Twitter.

Predictably, terms such as `britain`, `MPs`, `EU`, `brexit`, `leave`, `arron banks` and `deal` were among the most frequent terms in our experimental corpus. These terms are depicted in Figure 2 using a *word cloud* produced by *WordClouds.com* [32]. For clarity of purpose, we have also listed the 28 most frequent terms in our experimental corpus in Table 1.

To store the tweets that we collected, we saved them on a graph database using *Neo4j* [33]. Neo4j allowed us to analyse specific subsets of tweets separately. For example, Figure 3 displays a random sample of 3,411 users—represented by green circles—and 4,586 tweets—represented by blue circles. The green and blue circles are connected to indicate which specific user posted which specific tweet. The three red circles in Figure 3 represent hashtags: the one at the top is `#stopbrexit` and the one at the bottom is `#brexit`. The third hashtag in Figure 3 is `#brexitdeal`. Although we specifically retrieved tweets labelled with the hashtags `#brexit` and `#stopbrexit`, some tweets were labelled with additional hashtags. The most common hashtags in our experimental corpus were `#brexitdeal`, `#europe`, `#news`, `#politico`, `#politics` and `#telegraph`.

Figure 2. Experimental corpus word cloud

Table 1. Heading and text fonts.

Term	Occurrences	Term	Occurrences	Term	Occurrences	Term	Occurrences
MPs	1,840	week	1,754	make	1,625	sign	1,538
save	1,482	madness	1,477	eu	1,453	brexit	1,059
back	1,052	business	1,030	owner	1,022	tears	1,021
british	1,003	people	903	worried	878	fighting	866
staff	864	livelihood	860	upsides	856	leave	611
arron banks	507	deal	496	uk	457	world	395
theresa may	385	vote	326	government	295	today	287

As shown in Figure 3, the hashtag `#brexit` was far more common in our experimental corpus than `#stopbrexit`. Indeed, the total number of tweets labelled with the hashtag `#stopbrexit` in our experimental corpus was only 5% of the total number of tweets labelled with `#brexit`.

3.1. Validation of the Experimental Corpus

We needed to confirm if the topics that we discovered were indeed trending. To do so, we retrieved a collection of breaking news headlines published at around the same days we collected the tweets. These were the news headlines for 7th-10th March 2019. If our trends matched the news headlines on those days, we concluded we had discovered actual trends—for example, if one of our trending topics was identified and tweeted about by a well-regarded news source, such as the *BBC News* [34], we assumed we had accurately found a trending topic.

We collected breaking news headlines by means of the *News API* [35], an open source API that returns JSON metadata for news headlines and articles.

While other APIs, such as the *New York Times API* [36], provide information about news published by a single source, the News API offers information gathered from more than 30,000 sources worldwide. When we retrieved our experimental corpus, the News API supplied headlines published by 47 news sources in English language. Such sources, and the specific number of headlines published by each of them during the 24 hours that our retrieval of tweets lasted, are listed in Figure 4—note that the top 9 sources involved in our experiment published, each of them, 20 news headlines.

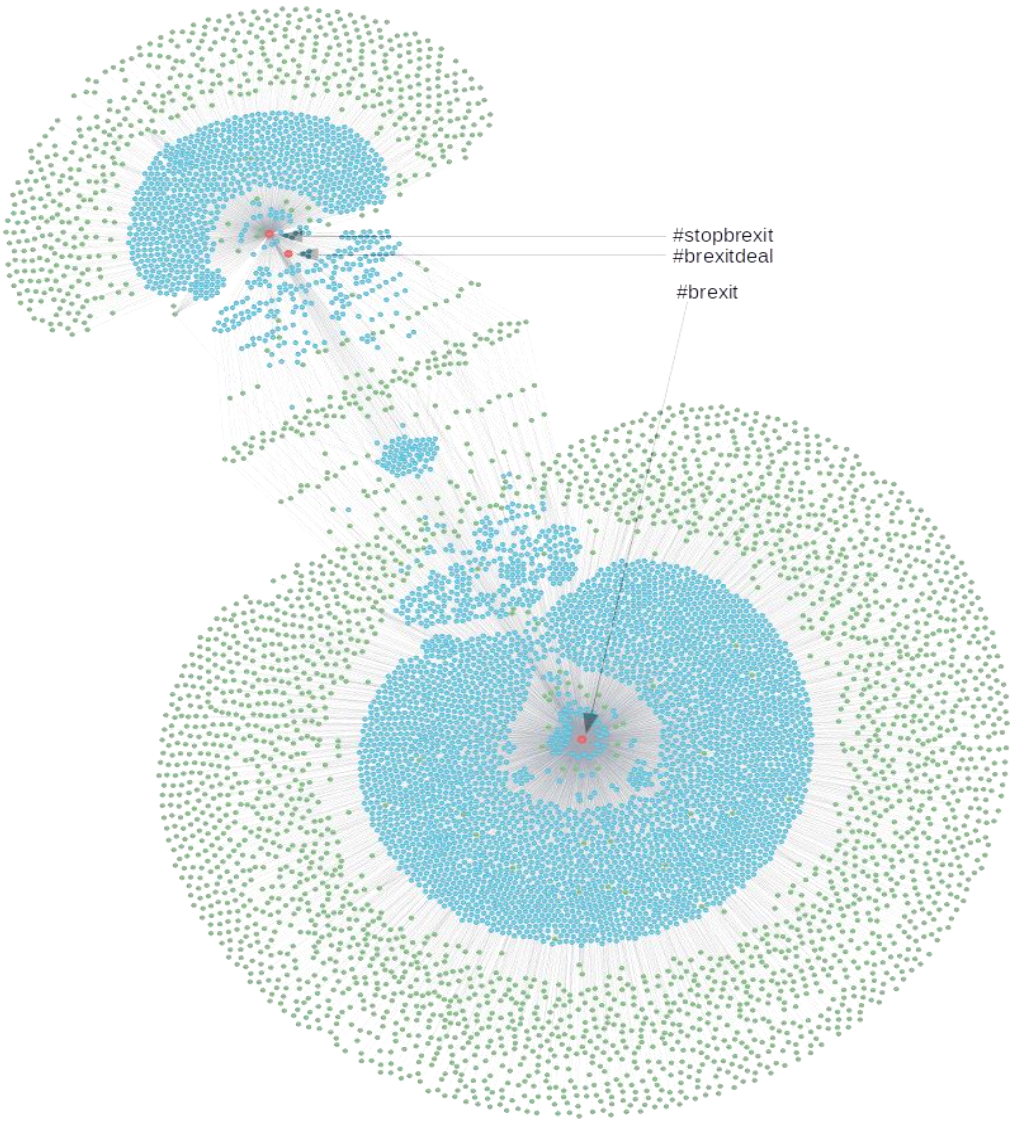

Figure 3. Graph database sample

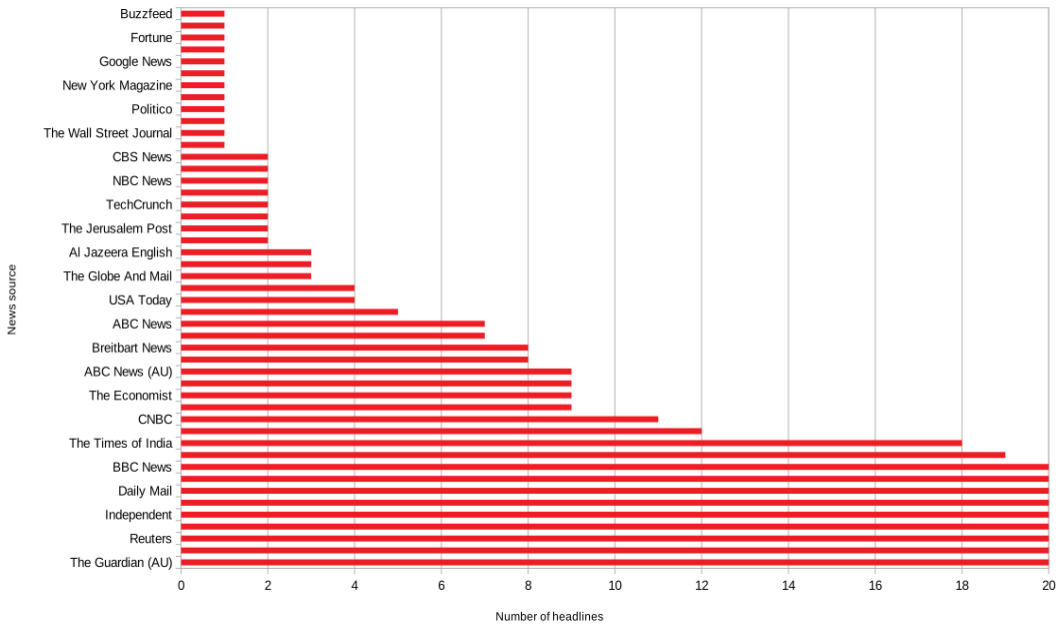

Figure 4. News API sources

4. TRENDING ALGORITHM

The definition of *trending topic* that we used is a modified version of the one proposed by Benhardus and Kalita [15]: *a trending topic is a word, or combination of words, that experiences an increase in usage, both in relation to its long-term usage and in relation to the usage of other words*. Our algorithm to identify trending topics is written in Python—see [37]. Figure 5 shows a diagram comprising the steps that we followed to identify trending topics. We will elaborate on the critical components of Figure 5 in the following subsections.

4.1. Tokenisation

Tokenisation is the process of splitting a piece of text into the different terms that compose it [38]. We tokenised tweets using Bonzanini's text pre-processing code [39], because it has become a popular tool for Twitter text mining.

4.2. N-Grams

An *n-gram* is a contiguous sequence of *n* terms from a piece of text. Such items can be phonemes, syllables, letters or words. For the purpose of our trending topic algorithm, we focused on *words*. We refer to each separate word as a *unigram*; two adjacent words as a *bigram*; three adjacent words as a *trigram*; four adjacent words as a *4-gram*; and so on. For example, consider the following tweet, posted on 8th March 2019 at 15:28:57, "`MPs could save Britain from Brexit`". Then, the n-grams are as follows:

- Unigrams: "`MPs`"; "`could`"; "`save`"; "`Britain`"; "`from`"; "`Brexit`";
- Bigrams: "`MPs could`"; "`could save`"; "`save Britain`"; "`Britain from`"; "`from Brexit`";
- Trigrams: "`MPs could save`"; "`could save Britain`"; "`save Britain from`"; "`Britain from Brexit`";

- 4-grams: “MPs could save Britain”; “could save Britain from”; “save Britain from Brexit”;

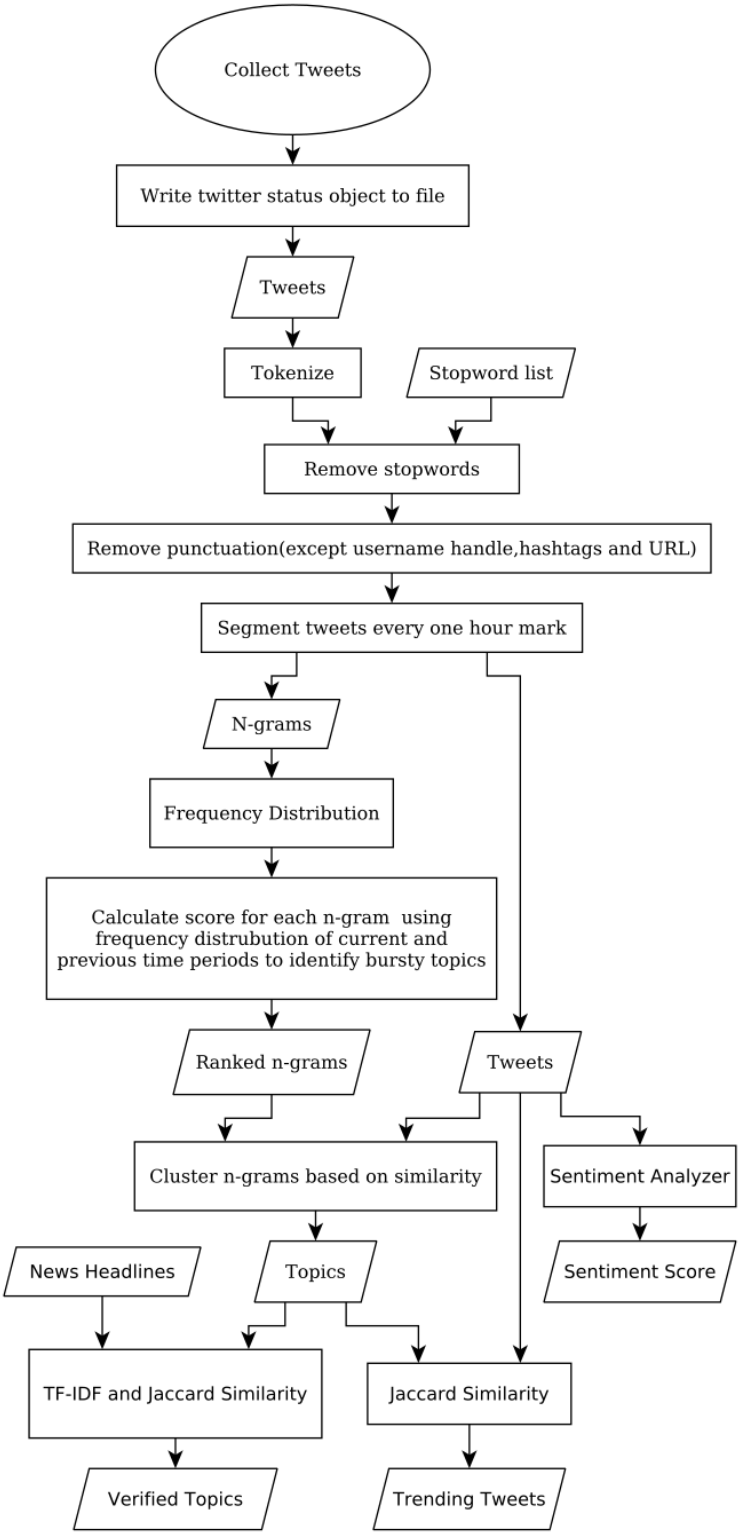

Figure 5. Determination of trending topics and sentiment analysis

Note that n-grams group together co-occurring terms. Using n-grams makes particular sense for Twitter, since a large number of tweets are just copies, or *retweets*, of previous tweets, so important n-grams tend to be frequent. To identify the most characteristic terms in our experimental corpus, we extracted all the unigrams, bigrams and trigrams, and they became our candidates to become trending topics.

4.3. Stop-words

Stop-words are extremely common and semantically non-selective words [40], such as *the*, *is*, *at*, *which* and *on*. By removing these words, we guarantee that our analysis concentrates on "meaningful" terms within tweets. To remove them, we used the stop-word list provided by Salton and Buckley for the experimental *SMART* information retrieval system [41], which contains 571 words. Given that tweets, as any other form of online text, usually contain some non-textual content [42]—unicode characters and punctuation—we also took advantage of this step to "clean up" the tweets.

4.4. Stemming

In our previous experiments—see Palomino et al. [43]—we *stemmed* the text of the tweets after removing stop-words. *Stemming* reduces inflectional and derivationally related forms of a word to a common base [38]. Stemming trims down the number of words in a corpus—and consequently the number of n-grams—by coalescing related terms into single stems—for example, the three words `presidency`, `president` and `presiding` become `presid`.

We expected stemming to increase the speed of our algorithm, as the number of n-grams to process would be smaller. However, our latest experiments show that the advantages of stemming are minimal. The number of bigrams and trigrams extracted per hour over a 24-hour period is, on average, 32,230 with stemming and 32,485 without stemming—these numbers refer to the n-grams derived from the tweets every hour for 24 continuous hours. Overall, processing 285 additional n-grams per hour is trivial. Hence, we have decided to omit stemming, though we are fully aware of the importance of pre-processing tweets [42], and we will continue to experiment until reaching definitive conclusions.

4.5. Trending Topic Identification

To identify trends, we extracted all the bigrams and trigrams in the corpus every hour. We avoided unigrams, because they are often too limited to characterise a topic [14]. We could have extracted bigrams and trigrams every 30 minutes to identify trends in shorter intervals, as we did before [43]. However, to test our algorithm, we worked on an hourly basis.

To discover trends as they surface, we took into account the changing frequency of n-grams over time. This uncovers emerging trends by comparing the frequencies of the n-grams in the current timeslot with those of preceding timeslots; thus distinguishing between trends that crop up in the past—for example, in the previous hour—but remain popular currently—as opposed to totally new trends that are surfacing right now.

We then proceeded to rank every bigram and trigram in the corpus on a per-hour basis using the df - idf_t metric defined by Aiello et al. [14] through the following formula:

$$df - idf_t = \frac{df_i + 1}{\log(\frac{\sum_{j=1}^{t} df_{i-j}}{t} + 1) + 1}$$

The formula stated above is computed for each bigram and trigram in the current timeslot—namely, *i*—based on its current *document frequency*—namely df_i. The current document frequency is the number of times the n-gram appears in the tweets posted during the current timeslot. The logarithm of the average of the n-gram's document frequencies in the previous *t* timeslots "downgrades" the weight of n-grams that are not new.

The top 100 bigrams and trigrams per hour, ranked according to the formula stated above, become our candidates to be trending topics. We arbitrarily chose the top 100, because it would have been computationally unfeasible to process more than that. Processing 100 n-grams on an Intel(R) Core(TM) i7-3517U CPU @ 1.90GHz requires about an hour, which is as much time as we have to identify the trends of the previous hour before processing the tweets for the following hour.

4.6. Jaccard Similarity Index

As indicated in Figure 5, once we rank the n-grams, we use a clustering algorithm to group the most representative n-grams into categories. This is because a single n-gram may not be enough to describe a trending topic, but a group of them typically offers relevant details of the topic.
The similarity measure we employ to cluster the n-grams is based on the *Jaccard similarity index* [44], which measures the "proximity" between two attributes by taking the intersection of both and dividing it by their union [45]. Thus, the similarity between two n-grams—namely, $n\text{-}gram_1$ and $n\text{-}gram_2$—is the number of tweets containing $n\text{-}gram_1$ and $n\text{-}gram_2$ simultaneously, divided by the number of tweets containing only $n\text{-}gram_1$ plus the number of tweets containing only $n\text{-}gram_2$. We assume that pairs of n-grams whose Jaccard similarity index is smaller than 0.2 represent the same topic and therefore can be clustered together.

To illustrate the identification of topics, Table 2 displays two of the highest ranked bigrams in our corpus between 5:28 and 6:28 on 9th March 2019—these bigrams correspond to `mark francois` and `british people`. Table 2 lists some of the tweets that were clustered together under these bigrams using Jaccard's similarity. Table 2 also indicates the number of times that each of these tweets was retweeted.

It should be observed that 8th March 2019, when we started to collect our experimental corpus, was precisely the day when the BBC aired an argument between the author Will Self and the Brexiteer MP Mark Francois on *Politics Live*, a popular television political programme. Our first trend in Table 2 captured the polemic derived from such a programme.

Although we do not have any formal evidence that Jaccard's similarity is better than other similarity measures, such as those discussed in [44], for the purpose of trend discovery, empirical tests indicate that Jaccard is a reasonable option.

4.7. TF-IDF

To verify if the topics which we discovered were indeed trending, we calculated the distance between our topics and the news headlines we captured using the *cosine similarity* [38]. As explained above, we looked for trending topics on 8th-9th March 2019, but we retrieved news

headlines for 7th-10th March 2019. This is because we realised that some of the trends which we discovered referred to headlines published the day before.

Some of the trends which we discovered near the end of our 24-hour experiment, appeared in the news the following day. Thus, we had to consider a larger range for the news headlines than for the experimental corpus.

Table 2. Examples of trending topics and some associated tweets.

Topic 1:	mark francois
Tweets:	(4x)>> RT @fractallogic1: Mark Francois is not a clown \smiley{} he's dangerous liar doing great harm to this country. #Brexit https://t.co/EpQcseorAm
	(2x)>> RT @BBCPolitics: "I think you should apologise.""To who? Racists and anti-Semites?"That stare Conservative MP Mark Francois and auth...
	(1x)>> RT @JerryHicksUnite: Who's `harder' Will Self or Mark Francois? There's only one way to find out FIGHT!#willself #MarkFrancois #politicsli...
	(1x)>> RT @StrongerStabler: Ultra-Brexiter Mark Francois is on #Peston to spout Brexit tripe. He is the Deputy Chair of the ERG \& its biggest cont...
	(1x)>> RT @ArgyleLoz: Will Self completely destroyed not just Mark Francois today, but Brexit as a whole, I've tried to engage with leave voters, ...
Topic 2:	british people
Tweets:	(13x)>> RT @Mike_Fabricant: If Parliament and this Government betray the clear will of the British people by either delivering no #Brexit or a half...
	(10x)>> RT @ajmpolite: May tells the world that #Brexit 'belongs to the British people' and 'everyone now wants to get it done'.Well, it doesn't b...
	(4x)>> RT @RedHotSquirrel: From a Government that does NOT WANT to Leave... For how much longer are we going to stomach this outrage? The British...
	(3x)>> RT @georgegalloway: ``This coming week will decide whether the democratic decision by the British people to leave the EU is honoured or not...
	(1x)>> Why people still listening to that lair Not one of he's devastating false predictions have come through He’s a fals... https://t.co/sxEHcsObpC
	(1x)>> RT @jkenney: Congratulations to the British people on choosing hope over fear by embracing a confident, sovereign future, open to the world...
	(1x)>> RT @Devon4Europe: @theresa_may How dare you say the British People want this settled without adding `and they now want to Remain, having se...

To compute the cosine similarity, we calculated the TF-IDF weights of the different terms in the headlines and our trending topics. TF-IDF is based on a composite of the *term frequency* and the *inverse document frequency* [46]. *Term frequency* can be defined as $tf_{ij} = n_{ij}$, where $n_{i,j}$ is the number of times word i occurs in tweet j and $\sum_k n_{kj}$ is the total number of words in tweet j. *Inverse document frequency* is defined as

$$idf_i = \log \frac{T}{d_i}$$

where d_i is the number of tweets that contain word i and T is the total number of tweets—namely, 62,397.

If the cosine similarity between a particular topic and a particular headline was smaller than 0.25, we assumed that we actually discovered a trending topic. Figure 6 plots the number of trending topics that we identified on an hourly basis, and examples of such topics. In total, our algorithm produced 147 trending topics over 24 hours. However, only 53 of them matched the news headlines—those are the ones plotted in Figure 6.

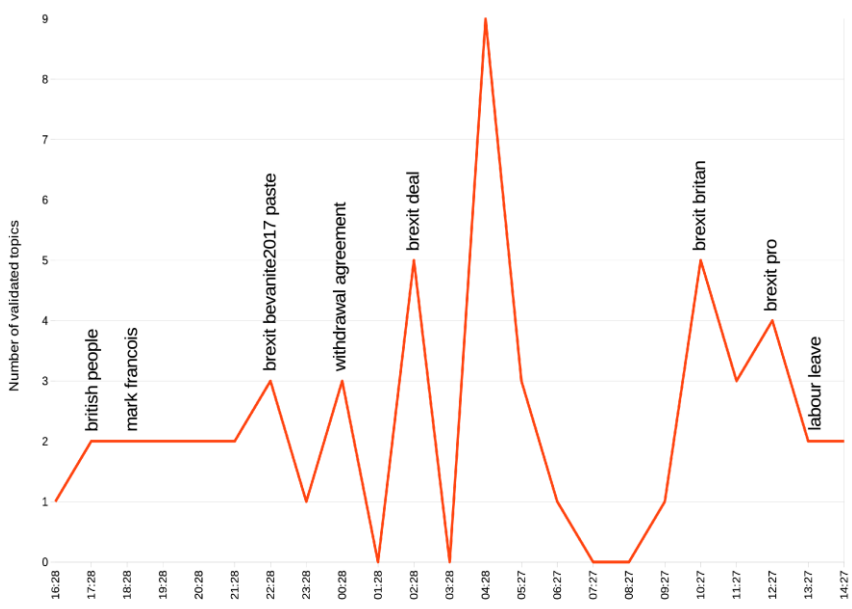

Figure 6. Validated trending topics

4.8. SentiStrength

To assess the impact of sentiment on trend discovery and exploration, we chose *SentiStrength* [24], a tool to determine the sentiment expressed in tweets [47]. Unlike other sentiment analysis tools, which tend to be commercially-oriented, SentiStrength was created by academics to exploit the de-facto grammars and spelling styles of the casual communication that is prevalent in social networks [48]. Figure 7 plots the number of positive, negative and neutral tweets in the entire corpus, according to SentiStrength. As readers can see, negative tweets—27,932 in total—are more common than neutral ones—23,377—and positive ones—11,088. Thus, we can confirm that the corpus is largely negative.

4.9. VADER

To prevent any biases derived from choosing a single sentiment analysis tool, we also used *VADER—Valence Aware Dictionary and sEntiment Reasoner*—a rule-based tool to identify sentiments expressed in social media [25]. We selected VADER because it is well-regarded among the academic community, and easy to adapt for research testing [26]. Figure 7 also plots

the number of positive, negative and neutral tweets in the corpus according to VADER. As readers can see, negative tweets—22,862—are more common than positive ones—22,347—and neutral ones—17,188.

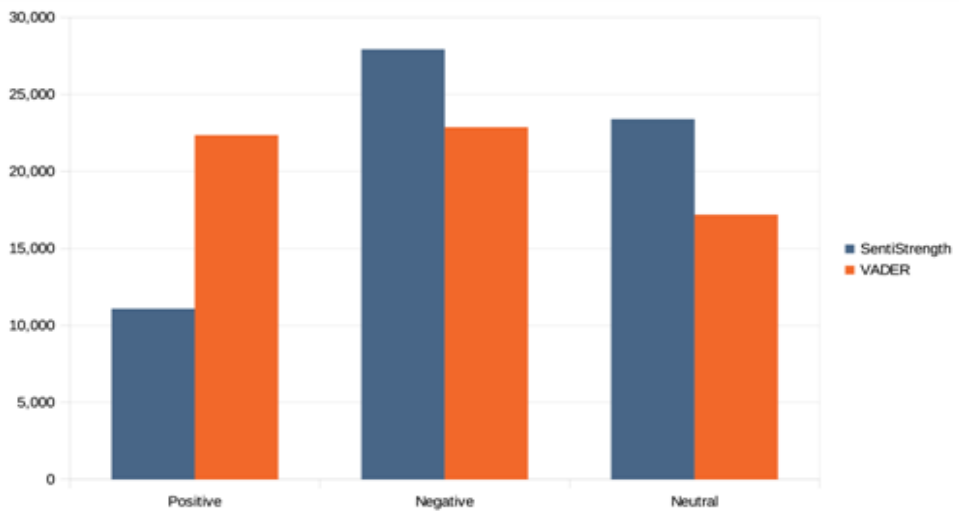

Figure 7. Validated trending topics

5. EXPERIMENTAL RESULTS

Researchers have proposed a number of metrics to estimate the overall sentiment expressed towards particular topics on social networks. A common metric for this purpose is the *net sentiment rate* (NSR) [49].

The NSR is defined as the subtraction of the number of negative conversations—negative tweets in our case—from the number of positive conversations—positive tweets—divided by the total number of conversations—total number of tweets. According to SentiStrength, the NSR of our experimental corpus is negative—to be precise, it is -0.27. According to VADER, the NSR of our experimental corpus is also negative—namely, -0.008. However, the NSR according to VADER is not as negative as SentiStrength estimates. This evidences the need for evaluating more than one tool, rather than relying on any one of them without further analysis.

Figure 8 and Figure 9 display how sentiment evolved over the 24-hour length of our experiment, according to SentiStrength and VADER, respectively. In the case of SentiStrength, the number of neutral tweets was consistently above the number of positive tweets for 24 hours; yet, the number of neutral tweets never exceeded the number of negative tweets. In the case of VADER, the number of neutral tweets was mostly below the number of positive tweets for the entire length of the experiment. Once again, the results demonstrate the lack of consensus among sentiment analysis tools.

6. CONCLUSIONS

It has been assumed that the emergence of trends in Twitter depends on only two factors: the influence of the users who are involved in the trend, and the adoption by the users who are exposed to the trend. We have investigated if trends are also related to the sentiment expressed in the tweets that form an emerging trend.

While we have produced preliminary results concerning the impact of sentiment on trends, further work is necessary to measure how influential highly polarised tweets are. We have made our code available online to other researchers who wish to reproduce our experiments. We expect to continue our investigation and include other domains of information. Although focusing on news headlines was a convenient starting point, other domains, such as sports and entertainment, might provide further insights.

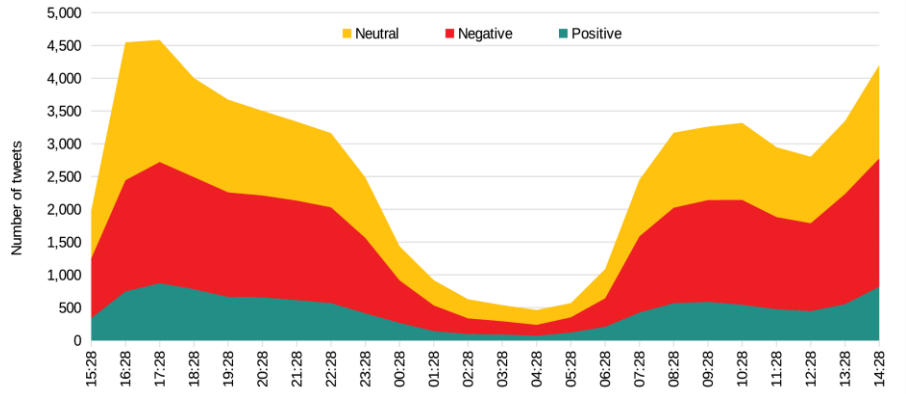

Figure 8. Sentiment of tweets per hour according to SentiStrength

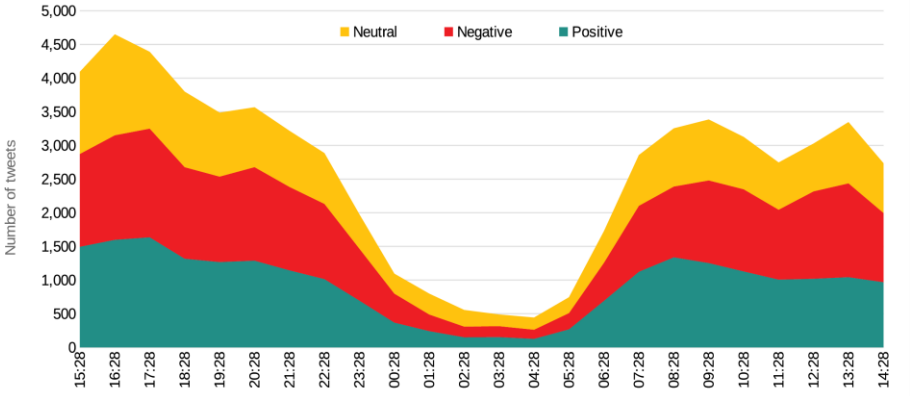

Figure 9. Sentiment of tweets per hour according to VADER

ACKNOWLEDGEMENTS

Marco Palomino gratefully acknowledges the funding provided by the Interreg 2 Seas Mers Zeeën AGE'IN project (2S05-014) to support his work in the research described in this publication. The authors are thankful to Martin Lavelle for reading the manuscript and providing insightful comments.

REFERENCES

[1] D. Boyd, S. Golder, and G. Lotan, "Tweet, Tweet, Retweet: Conversational Aspects of Retweeting on Twitter," in Proceedings of the International Conference on System Sciences (HICSS). Honolulu, HI:IEEE, Mar. 2010, pp. 1–10.

[2] S. Milstein and T. O'Reilly, The Twitter Book. O'Reilly Media, May 2009.

[3] A. Pak and P. Paroubek, "Twitter as a Corpus for Sentiment Analysis and Opinion Mining," in Language Resources and Evaluation Conference, vol. 10, no. 2010, 2010, pp. 1320–1326.

[4] G. D. Bodie and M. J. Dutta, “Understanding Health Literacy for Strategic Health Marketing: eHealth Literacy, Health Disparities, and the Digital Divide,” Health Marketing Quarterly, vol. 25, no. 1-2, pp. 175–203, 2008.
[5] N. Oliveira, P. Cortez, and N. Areal, “The Impact of Microblogging Data for Stock Market Prediction: Using Twitter to Predict Returns, Volatility, Trading Volume and Survey Sentiment Indices,” Expert Systems with Applications, vol. 73, pp. 125–144, 2017.
[6] Internet Live Stats. (2019) Twitter Usage Statistics. [Online]. Available: http://www.internetlivestats.com/twitter-statistics/
[7] C. Smith. (2019) 400 Interesting Twitter Stats and Facts (2019) by the Numbers. [Online]. Available: https://expandedramblings.com/index.php/twitter-stats-facts/
[8] B. R. Heravi and N. Harrower, “Twitter Journalism in Ireland: Sourcing and Trust in the Age of Social Media,” Information, Communication & Society, vol. 19, no. 9, pp. 1194–1213, 2016.
[9] N. Shahid, M. U. Ilyas, J. S. Alowibdi, and N. R. Aljohani, “Word Cloud Segmentation for Simplified Exploration of Trending Topics on Twitter,” IET Software, vol. 11, no. 5, pp. 214–220, 2017.
[10] Twitter, Inc. (2019) Twitter Trends FAQs. [Online]. Available: https://help.twitter.com/en/using-twitter/twitter-trending-faqs
[11] S. B. Hobolt, “The Brexit Vote: A Divided Nation, A Divided Continent,” Journal of European Public Policy, vol. 23, no. 9, pp. 1259–1277, 2016.
[12] BigCommerce Pty. Ltd. (2019) What is a Trending Topic and How can It Be Used In Ecommerce? [Online]. Available: https://www.bigcommerce.com/ecommerce-answers/what-is-trending-topic-ecommerce/
[13] C. Dewey, “Why Didn’t #FreddieGray Trend on Twitter?” The Washington Post, Apr. 2015, https://www.washingtonpost.com/news/the-intersect/wp/2015/04/27/why-didnt-freddiegray-trend-on-twitter/.
[14] L. M. Aiello, G. Petkos, C. Martin, D. Corney, S. Papadopoulos, R. Skraba, A. G¨oker, I. Kompatsiaris, and A. Jaimes, “Sensing Trending Topics in Twitter,” IEEE Transactions on Multimedia, vol. 15, no. 6, pp. 1268–1282, 2013.
[15] J. Benhardus and J. Kalita, “Streaming Trend Detection in Twitter,” International Journal of Web Based Communities, vol. 9, no. 1, pp. 122–139, 2013.
[16] S. Petrovi´c, M. Osborne, and V. Lavrenko, “Streaming First Story Detection with Application to Twitter,” in Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics. Association for Computational Linguistics, 2010, pp. 181–189.
[17] ——, “Using Paraphrases for Improving First Story Detection in News and Twitter,” in Proceedings of the Annual Conference of the North American Chapter of the Association for Computational Linguisticss: Human Language Technologies. Association for Computational Linguistics, 2012, pp. 338–346.
[18] D. A. Shamma, L. Kennedy, and E. F. Churchill, “Peaks and Persistence: Modeling the Shape of Microblog Conversations,” in Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work. ACM, 2011, pp. 355–358.
[19] D. Hiemstra, “A Probabilistic Justification for Using TF_IDF Term Weighting in Information Retrieval,” International Journal on Digital Libraries, vol. 3, no. 2, pp. 131–139, 2000.
[20] B. Pang and L. Lee, “Opinion Mining and Sentiment Analysis,” Foundations and Trends in Information Retrieval, vol. 2, no. 1–2, pp. 1–135, 2008.
[21] A. Giachanou and F. Crestani, “Like It or Not: A Survey of Twitter Sentiment Analysis Methods,” ACM Computing Surveys (CSUR), vol. 49, no. 2, p. 28, 2016.
[22] W. Medhat, A. Hassan, and H. Korashy, “Sentiment Analysis Algorithms and Applications: A Survey,” Ain Shams Engineering Journal, vol. 5, no. 4, pp. 1093–1113, 2014.
[23] F. N. Ribeiro, M. Ara´ujo, P. Gonc¸alves, M. A. Gonc¸alves, and F. Benevenuto, “Sentibench–A Benchmark Comparison of State-of-the-Practice Sentiment Analysis Methods,” EPJ Data Science, vol. 5, no. 1, p. 23, 2016.
[24] M. Thelwall. (2019) SentiStrength. [Online]. Available: http://sentistrength.wlv.ac.uk/
[25] C. J. Hutto and E. Gilbert, “Vader: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text,” in Proceedings of the International Conference on Weblogs and Social Media (ICWSM-14). Ann Arbor, MI: The AAAI Press, 2014.

[26] A. Connelly, V. Kuri, and M. Palomino, “Lack of Consensus among Sentiment Analysis Tools: A Suitability Study for SME Firms,” in Proceedings of the 8th Language & Technology Conference, Poznan, Poland, Nov. 2017, pp. 54–58.
[27] Tweepy. (2019) Tweepy Documentation. [Online]. Available: http://www.tweepy.org/
[28] Twitter, Inc. (2019) Twitter Developer Platform. [Online]. Available: https://developer.twitter.com/en/docs.html
[29] Tweepy. (2019) Streaming With Tweepy. [Online]. Available: http://docs.tweepy.org/en/v3.4.0/streaming how to.html
[30] Cambridge Network Limited. (2019) Best times to post on social media. [Online]. Available: https://www.cambridgenetwork.co.uk/news/best-times-to-post-on-social-media/
[31] J. Watts. (2019) Brexit vote: Theresa May says ‘we may never leave the EU’ if MPs reject her deal. The Independent. [Online]. Available: https://www.independent.co.uk/news/uk/politics/brexit-vote-theresa-may-deal-commons-speech-parliament-a8813861.html
[32] Zygomatic. (2019) Free Online Word Cloud Generator and Tag Cloud Creator - WordClouds. [Online]. Available: https://www.wordclouds.com/
[33] J. J. Miller, “Graph Database Applications and Concepts with Neo4j,” in Proceedings of the Southern Association for Information Systems Conference, vol. 2324, Atlanta, GA, 2013, p. 36.
[34] BBC. (2019) BBC News. [Online]. Available: http://www.bbc.co.uk/news
[35] News API. (2019) News API - Access Worldwide News with Code. [Online]. Available: https://newsapi.org/
[36] NYTimes.com. (2019) The New York Times Developer Network. [Online]. Available: https://developer.nytimes.com/
[37] A. Murali. (2019) Trending Topics. [Online]. Available: https://github.com/adithya2208/Trending-Topics
[38] H. Sch¨utze, C. D. Manning, and P. Raghavan, Introduction to Information Retrieval. Cambridge University Press, 2008, vol. 39.
[39] M. Bonzanini, Mastering Social Media Mining with Python. Packt Publishing Ltd, 2016.
[40] W. J. Wilbur and K. Sirotkin, “The Automatic Identification of Stop Words,” Journal of information science, vol. 18, no. 1, pp. 45–55, 1992.
[41] G. Salton, “A New Comparison between Conventional Indexing (MEDLARS) and Automatic Text Processing (SMART),” Journal of the Association for Information Science and Technology, vol. 23, no. 2, pp. 75–84, 1972.
[42] E. Haddi, X. Liu, and Y. Shi, “The Role of Text Pre-Processing in Sentiment Analysis,” Procedia Computer Science, vol. 17, pp. 26–32, 2013.
[43] M. A. Palomino, Q. Ribac, and G. L. Masala, “The Nature of Twitter Trending Topics - Analysing Intrinsic Factors Associated with the Twitter Ecosystem,” in The 17Th International Conference on Intelligent Software Methodologies, Tools, and Techniques. Granada, Spain: IOS Press, Sep. 2018.
[44] A. Strehl, J. Ghosh, and R. Mooney, “Impact of Similarity Measures on Web-Page Clustering,” in Workshop on Artificial Intelligence for Web Search (AAAI 2000), vol. 58, 2000, p. 64.
[45] L. Zahrotun, “Comparison Jaccard Similarity, Cosine Similarity and Combined Both of the Data Clustering with Shared Nearest Neighbor Method,” Computer Engineering and Applications Journal, vol. 5, no. 1, pp. 11–18, 2016.
[46] G. Salton and D. Harman, Information Retrieval. John Wiley and Sons Ltd., 2003.
[47] M. Thelwall, K. Buckley, G. Paltoglou, D. Cai, and A. Kappas, “Sentiment Strength Detection in Short Informal Text,” Journal of the American Society for Information Science and Technology, vol. 61, no. 12, pp. 2544–2558, 2010.
[48] M. Thelwall, K. Buckley, and G. Paltoglou, “Sentiment in Twitter events,” Journal of the American Society for Information Science and Technology, vol. 62, no. 2, pp. 406–418, 2011.
[49] M. Palomino, T. Taylor, A. G¨oker, J. Isaacs, and S. Warber, “The Online Dissemination of Nature–Health Concepts: Lessons from Sentiment Analysis of Social Media Relating to ”Nature-Deficit Disorder”,” International Journal of Environmental Research and Public Health, vol. 13, no. 1, p. 142, 2016.
[50] Y. Zhang, X. Ruan, H. Wang, H. Wang, and S. He, “Twitter Trends Manipulation: A First Look Inside the Security of Twitter Trending,” IEEE Transactions on Information Forensics and Security, vol. 12, no. 1, pp. 144–156, 2017.