

Detection of Sentiment Provoking Events in Social Media

Hoda Daou

Olayan School of Business - American University of Beirut

hd66@aub.edu.lb

Abstract

Social media has become one of the main sources of news and events due to its ability to disseminate and propagate information very fast and to a large population. Social media platforms are widely accessible to the population making it difficult to extract relevant information from the huge amount of posted data. In our study, we propose an algorithm that automatically detects events using strong sentiment classification and appropriate clustering techniques. We focus our study on a specific type of events that triggers strong sentiment among the public. Results show that the suggested methodology is able to identify important events, such as a mass shooting and plane crash, using a generalized and simple approach.

1. Introduction

Advancement in technologies and data access capabilities have enabled both the development of online social media platforms and the birth of big data as a new field of study. Social media sites such as Facebook, Twitter and Instagram have played a big role in changing the way individuals communicate and even the way businesses interact and perform [7]. Whether in marketing, sales or public relations social media analytics has proven to be very valuable and if used properly with the right business goals and quantitative metrics, can provide great insights and even increase reach, awareness and purchase intent [15].

Social media analytics has become a very active research area mainly due to the availability of the data, the novel analytics algorithms and its application across different domains [15, ch.1]. Whether in marketing, sales or public relations social media has proven to be very valuable and if used properly with the right business goals and quantitative metrics, can provide great insights and even increase reach, awareness and purchase intent [15].

Events shared on social media are considered to be

real world occurrences that are localized in time and space and can range from social events to major natural disasters. Automatic detection of these events helps in identifying emerging topics and in monitoring and summarizing the information on social media. It also provides the potential for understanding and describing real actual events and improving the quality of higher level applications in the fields of social media analytics and traditional news detection.

Events in social media can range from social events such as concerts to a major natural disaster such as earthquake and even political events such as strikes that can cause major disruptions. In general, events are real world occurrences that are localized in time and space [7]. Event detection is the problem of automatically identifying significant incidents by analyzing social media data [19]. This task is considered to be very complex as it encompasses many different disciplines and the nature of the data used in the scope of this research, social media data, adds to its complexity.

Event detection from social media sources poses new issues and challenges mainly due to the data being analyzed. Data can be unstructured and is written by anyone and therefore can include informal, irregular words with spelling and grammatical errors [4]. For example, in the case of Twitter, streams sometimes contain meaningless messages [16], polluted content [18] and inaccurate information [12]. Preparing the ingested data involves complex analysis in order to reach reliable and clean data ready to be processed for building features, extracting events and even categorizing identified events.

This work focuses on detecting events that can trigger extreme sentiment among the population, reflected in social media posts. Using supervised classification, we show that real events can be detected from sentiment strength analysis applied on real posts from twitter.

This paper is organized as follows. The first section presents major previous work done in event detection. The methodology used in terms of pre-processing the

data, computed features and description of the different datasets used, is described in section 3. Section 4 focuses on choosing the right sentiment classification algorithm, which will be used as the first step in event detection, presented in section 5. We end this paper with discussion, conclusion and suggestions for future work.

2. Related Work

Research in event detection can be seen as similar to topic detection and tracking (TDT), Information Retrieval (IR) [4] [19] and topic modeling [3, 8]. TDT aims at finding events, detecting first stories which is feasible using tracking techniques and similarity comparison [6] [5], and following and tracking events in a stream of broadcasts [4].

Different techniques have been used in the literature for event detection. TF-IDF is characterized to be the dominant technique in text mining used for document representation and term weighting, more specifically for new event detection. It is also used by Brants *et.al* in an incremental model that uses new vocabulary as more stories are read and assigns weights according to their occurrence in the new documents [10]. Their goal is to build a system that detects stories about previously unseen events in a stream of news stories. Only the first story of an event is to be flagged and thus avoiding redundancy. Normalized term weights are computed to calculate the similarity between documents, which is used to score stories and identify new events that are higher than a certain threshold.

Event detection in social media is an important research area, widely seen in the literature. In [14], events posted in social media platforms are defined as real world events that occur within similar time periods and geographical locations, and are presented in different forms such as images, videos or texts. The paper focuses on the multi-scale characteristic of events: temporal and spatial scales. Different types of events can span over different temporal and geographical scales. Authors analyze the relationship between the temporal and spatial scales in order to separate events that extend over different scales. Similarity scores between pairs of tweets that examine both temporal and spatial scales are used for that purpose. These scores are used to build a network graph of tweets as nodes and weighted edges based on the value of the similarity score. Clusters of events of different temporal and spatial scales are then detected from this similarity graph.

Tweets are also used as data source for event detection in [22], where the authors aim at handling issues of noise and scalability commonly faced in

social media data. Event Detection with Clustering of Wavelet-based Signals, EDCoW, is a model suggested by the authors that first uses individual words to construct signals that captures bursts of words, which is a conventional indication of the onset of an event. Wavelet transforms are used to compute these signals and filter out trivial words based on signal auto-correlation. Cross-correlation between signals is also computed to measure similarity between words which are represented by the constructed signals. This pair-wise cross correlation matrix is taken as the adjacency matrix of a sparse undirected weighted graph where the vertices are the signals and the edges weights are the cross correlation values. Event detection becomes possible with graph partitioning where each subgraph corresponds to an event that contains words with high cross correlation; and different events correspond to different partitions where the cross correlation between words is low.

Charu *et.al* [4] describe the key challenges for event detection in social streams as incorporating both the graphical structure that captures interactions and the temporal information that captures bursts, and handling large volumes of posts. In this work, a social stream is characterized as a continuous and temporal sequence of objects represented by its content, which is mostly text, the originating node and set of receiver nodes. Content and linkage information are used to dynamically create clusters of these objects. These clusters keep changing, either by assigning more points to a cluster (evolution of an existing event) or by adding new clusters (novel event), as the data stream evolves with time. For this purpose, the authors propose an online partition-based clustering methodology that keeps updating a set of clusters along with their summaries, as new social stream objects arrive. This update is done based on structural and content similarity measures between the new object and the clusters. Changes to these clusters are used to flag new events.

Graphical structure is also explored in [20] where the authors target episodes which are events that are part of a large context known as saga. The proposed algorithm constructs a keyGraph where nodes are keywords which are extracted terms from the text data and edges are formed when terms co-occur in a document. Document Frequency is computed for the extracted terms and keywords with low values are removed. Edges with low co-occurrence, lower than a pre-determined threshold, are also removed. With the assumption that keywords co-occur when there is relevant relationships, the authors use betweenness centrality to remove edges with high scores and therefore extracting communities. Since keywords can appear in more than one event, the

algorithm allows duplication of nodes for cases of high conditional probability of edges. Data from Live Labs' Social platform and blog posts are used in testing.

Event detection is applied on slightly different type of data where the authors aim at extracting real world events from Flickr images [13]. They use images, with user-supplied tags and other metadata such as time and location coordinates to group images to different events, as input data. The authors first map each tag to a 3D point composed of two dimensional location and time as third dimension. Wavelet transform is then used to decrease noise and provide multi-resolution analysis of tag distributions. This transform is able to emphasize regions where points are more dense and thus helps dismiss weak information. This method allows the distinction between aperiodic and periodic events where photos are taken at a sequence of time points with equal intervals. It also allows the analysis of the temporal and geographical distributions of tag occurrences in order to discover event-related tags with significant distribution patterns such as bursts in both dimensions. Clustering groups event-related tags into clusters; each representing an event with similar temporal and geographical distribution patterns as well as similar associated photos.

James Alan *et.al* [6] focus on new events and describe their detection as identifying stories that have not been reported before. They use a stream of broadcast news and try to determine relationships between stories that discuss real-world events. Each document is represented by a set of belief values calculated based on term frequency and inverse document frequency. The system first selects N most frequent features, then assigns a score to the text by using the #WSUM operator of the retrieval engine INQUERY for comparing a document to a query [11]. For that purpose, the *eval* metric is computed, that is the sum of the weighted beliefs divided by the sum of the weights. Belief values are produced by INQUERY's belief function using term frequency and inverse document frequency. This metric assigns a score to the text. A threshold is then used to determine whether the text reflects the onset of a new event. Observations show that there are surprising features of new events that have not been mentioned a lot in the past. These features can be new places, people or a specific word such as death and earthquake. To measure this surprising values, the authors compute a metric called surprise that is based on the distance between the occurrence of a word and all past occurrences. In addition, the study shows that stories about the same event often occur in clumps, more specifically for unexpected events such as disasters and major crimes. We would like to note that such events are

slang phrase	actual phrase
*4u	,Kiss for you
10q	thank you
10x	thanks
143	I love you
1daful	wonderful
1sec	one second
2b	to be
2bh	to be honest
2da	to the
2day	today

Table 1. slang mapping examples

emoticon	polarity measure
\$:	-1.5
%)	-0.4
%-)	-1.5
&-:	-0.4
&:	-0.7
(%:	-0.9
(':-	2.2
((:-	2.1
(:-0	2.8
(-;	1.3

Table 2. VADER emoticons polarity

often accompanied with strong sentiment, which partly explains the surge of social media posts after the events' onset, and is the major motivation behind our paper.

3. Methodology

The study focuses on events that are shared on social media and trigger strong sentiment among the population. This source of data is characterized to contain informal and irregular text which makes it challenging to apply mining techniques [7]. To mitigate the effect of this, pre-processing of social media posts is essential. For that purpose, we replace well known slang phrases with exact words using the same mapping used in [2]. In addition, emoticons are replaced with their meaning or sentiment using the same mapping as in VADER lexicon [17], explained in Section 3.1. Examples of slang phrases and emoticons are shown in Tables 1 and 2.

3.1. Features

Structural and polarity-based features are used in our supervised model. The main feature used in the overall classification is the mean value of prior polarity

of all terms, taking into account synonymous terms and negation. These scores are calculated using python implementation of VADER (Valence Aware Dictionary for sEntiment Reasoning) which is a purely rule-based algorithm, suggested in [17]. The algorithm uses a gold standard sentiment lexicon that is specific to microblog-like context in social media. The use of punctuation, prepositions, and upper case letters in social media posts tend to indicate a certain level of sentiment. Therefore, additional features are used:

- Punctuation: number of punctuation in the whole text and the number of punctuation in upper case words as additional feature.
- Prepositions: number of prepositions.
- Upper Case Words: number of words written in upper case.
- Count of Negative Polarity terms: number of terms with negative polarity
- Count of Positive Polarity terms: number of terms with positive polarity

3.2. Data

Data for Sentiment Classification Model Training

The dataset used to train the sentiment classification model incorporates 9743 posts from various sources: myspace (1041), digg (1084), youtube (3407) and twitter (4242). Data is human labeled using two different tags [1, 21]:

1. positive: positive sentiment strength score
2. negative: negative sentiment strength score

Posts were collected in 2010 and 2011 with users located in the United States. When labeling the data, subjects were given verbal instructions for coding each text as well as a booklet, containing a list of emoticons and acronyms with explanations, that clearly explains the task [21]. In addition, since the judgments should give a consistent perspective on sentiment in the data and not an estimate of the population average perception, the authors also conduct testing to identify a homogeneous subset of subjects by removing the ones that give anomalous scores [21].

Sentiment scores for terms are in the range 2 to 5 (positive) or -2 to -5 (negative). A score of -1 means neutral. In our study we aim at building one model that will detect strong sentiment that could be either negative or positive. We first combine both labels, positive label L_p and negative label L_n , to one label L , by applying

Sentiment Score	Number of Samples
-5	71
-4	327
-3	1010
-2	2358
-1	2067
2	1536
3	1670
4	583
5	121

Table 3. Number of samples per class

this equation: $L = s \times \max(\text{abs}(L_p), \text{abs}(L_n))$, with $s = -1$ if the maximum term is L_n and 1 otherwise. Then we label strong scores of $[-5, -4, 4, 5]$ to 1 and moderate scores of $[-3, -2, -1, 2, 3]$ to 0. This mapping is done since our focus is only detecting strong sentiment, which is class 1 in this case. Our classification problem now targets a binary outcome.

In general when examining social media posts, strong sentiment occurs less frequently than moderate ones. In this study our focus is to build a supervised classification model. Supervised algorithms are trained using labeled data and thus are sensitive to imbalance. Table 3 shows the number of samples in each class in the original data, summing up to high imbalance of 1102 samples for class 1 and 8641 samples for class 0.

Data for Event Detection

Another dataset is used for running the trained classification model and extracting events. This data was captured by streaming Twitter for one week, interrupted by a few uncaught exceptions and unforeseen technical issues, starting from May 18th. During that week, a major shooting events and even a plane crash occurred and therefore, posts are rich with events that could trigger extreme sentiment. The data includes around 100,000 tweets, written in English.

4. Supervised Sentiment Classification

4.1. Algorithm

The focus of this study is to choose a proper supervised machine learning algorithm for classification of strong sentiment. First the issue of imbalance should be resolved in order to improve the learning of the supervised algorithm. We test four different options of handling imbalance: (1) random repetition of samples, (2) Synthetic Minority Over-sampling Technique (SMOTE) [9], (3) no over-sampling and (4) down-sampling the majority class.

By down-sampling, the same number of samples of the minority class, 1102 samples, are randomly selected from the majority class which includes scores -3 to $+3$. This random selection conserves the same distribution of these scores as in the original one. Regarding SMOTE, the algorithm over-samples the minority class by creating synthetic minority class examples using a join of the nearest neighbors [9].

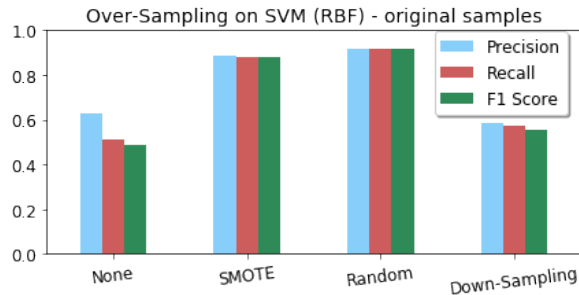


Figure 1. Test results of over-sampling techniques with Support Vector Machine - radial basis function (score computed on original samples only).

Several classification algorithms, such as logistic regression, nearest neighbors with $k = 3$, support vector machine (linear and radial basis function - SVM (RBF)) and naive bayes, were tested on the labeled dataset.

4.2. Testing & Results

Data used in training and testing includes both original and simulated samples. For choosing the proper over-sampling technique, training was done on 70% of the samples that are randomly selected, while testing was done on the remaining 30%. Sampling is done making sure original samples are included in both the training and testing sets. Testing showed that even though, SMOTE's complex algorithm might seem as a better candidate, simple random repetition had the same amount of improvement on classification compared to training on imbalanced original samples. This is highlighted in Figure 1 where SVM (RBF) algorithm is used and scores are computed only on original samples in the test set. This is done to ensure that results are not reflective of classification of simulated samples but of original ones.

When assessing classification, false positives, true positives and false negatives are computed for each class c . True positives (TP) are the number of correctly predicted positive samples, false positives (FP) are the number of times the actual class is not equal to c but predicted class is; and false negatives (FN) is the number of times the actual class is equal to c but

predicted class is not. The following metrics are used when assessing the results and comparing the different techniques: $Precision = \frac{TP}{TP+FP}$, $Recall = \frac{TP}{TP+FN}$ and $F1Score = 2 \times \frac{Recall \times Precision}{Recall + Precision}$.

For selecting the proper classification algorithm, we apply a 10-fold cross validation on the randomly oversampled dataset. Since oversampling created a big number of synthetic samples, we need to make sure performance scores are computed on original samples. For that purpose, we use the same test set for the different folds that is composed of original samples, and chosen across all scores of the different classes. We also make sure the training sets have a balanced dataset. Table 4 shows the mean values of the performance metrics across the 9 different training folds. Values show that SVM (RBF) algorithm gave the best results in terms of precision, recall and F1 scores.

5. Event Detection

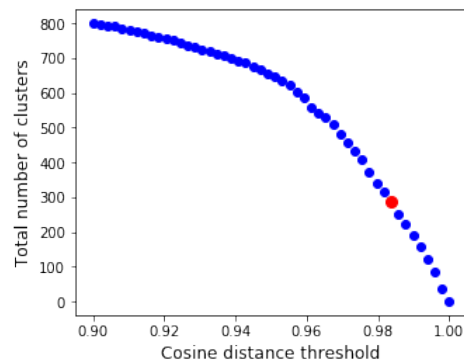


Figure 2. Results - Computing scores on all test samples.

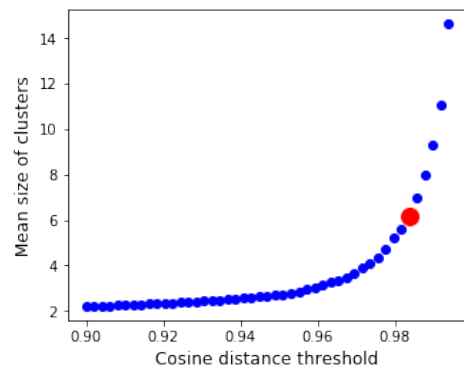


Figure 3. Results - Computing scores on all original samples.

Algorithm	Logistic Regression	Nearest Neighbors	Linear SVM	RBF SVM	Naive Bayes
Precision	0.578	0.667	0.582	0.796	0.591
Recall	0.672	0.849	0.668	0.914	0.564
F1-Score	0.566	0.688	0.581	0.840	0.573

Table 4. Performance Results with 10-fold cross-validation applied on oversampled data with random repetition.

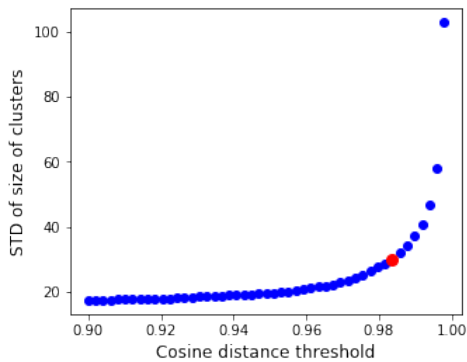


Figure 4. Results - Computing scores on all original samples.

To detect events, posts with extreme sentiments are first identified as having a strong classification probability. Then, in order to group together posts that discuss similar topics, we apply clustering on the Term Frequency - Inverse Document Frequency (TF-IDF) values of the content of these posts. Afterwards ranking the resulting events is done using a score that reflects the intensity of sentiment inside the cluster. These different steps used to identify the top events are described in detail in this Section.

As highlighted in the previous section, SVM (RBF) trained on labeled data showed good performance in classifying extreme sentiment in social media posts. This trained model is applied on the twitter unlabeled streamed data in order to detect the tweets that display strong sentiment, both negative and positive. Preliminary results show that a large portion of tweets was classified as having strong sentiment. This is due to the fact that the model was trained on data that might have different characteristics since it is coming from different social media platforms, and thus noise is to be expected. To solve that, we apply a threshold of 0.8 on the probability score of the classification output. Which means that only posts with probability greater than 0.8 are classified as displaying strong sentiment. Therefore,

we have stronger confidence in the outcome, and end up with 1162 strong sentiment-tweets, around 11% of the original set of tweets.

To better summarize the remaining tweets and extract relevant related events, we apply clustering to group the ones with similar characteristics. This is done based on Term Frequency - Inverse Document Frequency of the different words in the tweets. K-means and hierarchical clustering were tested and results showed that better results were obtained from hierarchical clustering with cosine distance. Choosing an appropriate value for the cosine distance threshold was done based on number of clusters, average cluster size and standard deviation of cluster size. Results are shown in Figures 2, 3 and 4, respectively. It should be noted that outliers having a very large number of tweets in the cluster were removed when computing the mean and standard deviation. The threshold is chosen in a way to have around 6 posts of strong sentiment, on average, in the different clusters, and a standard deviation of around 30. This chosen threshold is equal to 0.9836 and gives a total of 287 clusters. The clusters of events were ranked using a score equal to the ratio of number of posts having a sentiment score above 8.5 over the total number of posts in the cluster. The threshold used to determine strong sentiment posts is chosen after having tested several values between 8 and 9. From these 287 clusters, we show only the top 15 events that have displayed strong sentiment based on the scoring metric. We summarize the content of each cluster by displaying the top 7 words that occur in all posts belonging to that cluster. Table 5 shows the resulting top clusters that extend over a duration of 1 week. The algorithm was able to detect the main two events: Cuba plane crash and Texas school shooting and even rank them as having the highest scores. In addition it was able to detect a train derailment in Italy, a massacre in India, and important announcements that triggered strong social media reactions. It should be noted that from those 17 resulting clusters, one seems to contain miscellaneous tweets related to fake news, shown in the 9th row of

Table 5.

6. Discussion

The suggested methodology aims at detecting events that trigger strong sentiment such as disasters and important announcements. As first step, supervised classification algorithm is used to build a model that is able to extract posts with strong sentiment. Results show that with proper over-sampling and selection of the right machine learning technique, classification with reliable performance can be achieved. This is shown in Figure 1, 2 and 3 where random over-sampling and SVM (RBF) proved to be the best candidates among the ones tested in our study. As second step, the machine learning model is used to detect posts with strong sentiment from twitter starting May 18 2018 for a period of one week. Clustering with proper selection of distance threshold, shown in Figures 4, 5 and 6, is used to group posts of similar topics. As last step, selecting clusters of relevant size helped identifying the top 17 events that were strongly discussed in Twitter, shown in Table 4.

7. Conclusion & Future Work

The methodology summarized in the previous section proved to be successful in identifying important events. However improvements are still necessary, such as optimizing cosine distance threshold, probability score threshold when classifying sentiment and minimum cluster size in the last step. This can be achieved by using labeled data, which is an improvement the authors plan on focusing on next. In addition, the study disregards the time aspect of events. This is an important criteria that could be used in eliminating noise and achieving improvement in the clustering step. It could also supplement the methodology by being able to identify new events and even tracking their evolution in social media with time. This work showed that strong sentiment can help detect the onset of events. This will be tested and applied as general application to different social media platforms. In addition, the application of strong sentiment will be further investigated for other event detection algorithms to help identify events that are likely to be of interest.

8. References

- [1] Sentistrength, 2018. <http://sentistrength.wlv.ac.uk/>.
- [2] AGARWAL, A., XIE, B., VOVSHA, I., RAMBOW, O., AND PASSONNEAU, R. Sentiment analysis of twitter data. In *Proceedings of the Workshop on Languages in Social Media* (Stroudsburg, PA, USA, 2011), LSM '11, Association for Computational Linguistics, pp. 30–38.
- [3] AGGARWAL, C. C. *Data Streams - Models and Algorithms*. Springer, 2007.
- [4] AGGARWAL, C. C., AND SUBBIAN, K. Event detection in social streams. In *SDM (2012)*, SIAM / Omnipress, pp. 624–635.
- [5] ALLAN, J., LAVRENKO, V., AND JIN, H. First story detection in tdt is hard. In *Proceedings of the Ninth International Conference on Information and Knowledge Management* (New York, NY, USA, 2000), CIKM '00, ACM, pp. 374–381.
- [6] ALLAN, J., PAPKA, R., AND LAVRENKO, V. On-line new event detection and tracking. In *Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (New York, NY, USA, 1998), SIGIR '98, ACM, pp. 37–45.
- [7] ATEFEH, F., AND KHREICH, W. A survey of techniques for event detection in twitter. *Comput. Intell.* 31, 1 (Feb. 2015), 132–164.
- [8] BLEI, D. M., AND LAFFERTY, J. D. Dynamic topic models. In *Proceedings of the 23rd International Conference on Machine Learning* (New York, NY, USA, 2006), ICML '06, ACM, pp. 113–120.
- [9] BOWYER, K. W., CHAWLA, N. V., HALL, L. O., AND KEGELMEYER, W. P. SMOTE: synthetic minority over-sampling technique. *CoRR abs/1106.1813* (2011).
- [10] BRANTS, T., CHEN, F., AND FARAHAT, A. A system for new event detection. In *Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Informaion Retrieval* (New York, NY, USA, 2003), SIGIR '03, ACM, pp. 330–337.
- [11] CALLAN, J. P., CROFT, W. B., AND HARDING, S. M. The inquiry retrieval system. In *DEXA* (1992).
- [12] CASTILLO, C., MENDOZA, M., AND POBLETE, B. Information credibility on twitter. In *Proceedings of the 20th International Conference on World Wide Web* (New York, NY, USA, 2011), WWW '11, ACM, pp. 675–684.

Rank	Cluster Size	Top Words	Date	Actual Event
1	11	killed, pakistani, shooting, family, texas, says, diplomat	May 20	Father of a Pakistani girl killed in a Texas school shooting makes a declaration
2	95	school, shooting, texas, high, dead, killed, found	May 18	Texas school shooting: explosive devices found
3	68	dead, crash, plane, mourn, cubans, bodies, officials	May 21	Cuba plane crash: officials ID 20 bodies
4	15	disgusting, country, sydney, 3novices, words, gold, government	May 24	Australia news: boy discovered a dead fly trapped inside a sealed Le Snak pack.
5	28	great, news, congratulations, congrats, chiefpalmer, chendriekx, many	May 23	Announcement of the extension to Chief Adam Palmers contract
6	21	paul, republican, rages, warfare, media, jeopardy, ryans	May 24	Announcement from media claiming that Paul Ryan's job in jeopardy as Republican warfare rages
7	18	kills, injures, train, italy, derailment, crossing, road	May 24	Italy rail crash
8	11	world, discovery, closure, soldiers, plane, aviation, russia	May 24	Discovery of World War II soldier's plane brings closure to a long search
9	14	fake, news, already, from, lost, failing, president	NA	Miscellaneous tweets
10	11	irish, jeremy, calls, government, amid, role, corbyn	May 24	Jeremy Corbyn wants the Irish government to have a role in Northern Ireland
11	12	cuba, begin, crash, deadly, funerals, search, clues	May 21	Cuba plane crash: funerals begin
12	11	global, breaking, today, threat, vexes, makers, forex	May 24	Donald Trumps threat on new import tariffs
13	92	another, guns, today, killed, about, shootatweet, houston	May 20	Texas school shooting
14	11	deaths, thoothukudi, govt, protesting, "dmks", against, tipperary	May 22	Thoothukudi massacre: paramilitary and Tamil Nadu Police forces fired into a crowd of 20,000 people.
15	491	leadership, summit, awards, business, winners, announced, european	May 24	Winners of the European Leadership Awards 2018 announced at the European Business Summit

Table 5. Top 15 Events Detected

- [13] CHEN, L., AND ROY, A. Event detection from flickr data through wavelet-based spatial analysis. In *Proceedings of the 18th ACM Conference on Information and Knowledge Management* (New York, NY, USA, 2009), CIKM '09, ACM, pp. 523–532.
- [14] DONG, X., MAVROEIDIS, D., CALABRESE, F., AND FROSSARD, P. Multiscale event detection in social media. *CoRR abs/1404.7048* (2014).
- [15] FINGER, L., AND DUTTA, S. *Ask, Measure, Learn: Using Social Media Analytics to Understand and Influence Customer Behavior*. O'Reilly Media, 2014.
- [16] HURLOCK, J., AND WILSON, M. L. Searching twitter: Separating the tweet from the chaff. In *ICWSM* (2011), L. A. Adamic, R. A. Baeza-Yates, and S. Counts, Eds., The AAAI Press.
- [17] HUTTO, C. J., AND GILBERT, E. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *ICWSM* (2014).
- [18] LEE, K., EOFF, B. D., AND CAVERLEE, J. Seven months with the devils: a long-term study of content polluters on twitter. In *In AAAI Intl Conference on Weblogs and Social Media (ICWSM)* (2011).
- [19] PANAGIOTOU, N., KATAKIS, I., AND GUNOPULOS, D. Detecting events in online social networks: Definitions, trends and challenges. In *Solving Large Scale Learning Tasks. Challenges and Algorithms - Essays Dedicated to Katharina Morik on the Occasion of Her 60th Birthday* (2016), pp. 42–84.
- [20] SAYYADI, H., HURST, M., AND MAYKOV, A. Event detection and tracking in social streams. In *In Proceedings of the International Conference on Weblogs and Social Media (ICWSM 2009)*. AAAI (2009).
- [21] THELWALL, M., BUCKLEY, K., PALTOGLOU, G., CAI, D., AND KAPPAS, A. Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology* 61, 12 (December 2010), 2544–2558.
- [22] WENG, J., YAO, Y., LEONARDI, E., AND LEE, F. Event detection in twitter.