ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/359392243

Event detection in twitter by deep learning classification and multi label clustering virtual backbone formation

Article *in* Evolutionary Intelligence · March 2022

DOI: 10.1007/s12065-021-00696-6

CITATIONS
0

READS
83

4 authors, including:

Zahra Rezaei
University of Kashan
9 PUBLICATIONS **112** CITATIONS

SEE PROFILE

Behnaz Eslami
Islamic Azad University Tehran Science and Research Branch
7 PUBLICATIONS **9** CITATIONS

SEE PROFILE

All content following this page was uploaded by Behnaz Eslami on 22 March 2022.

The user has requested enhancement of the downloaded file.

Evolutionary Intelligence
https://doi.org/10.1007/s12065-021-00696-6

RESEARCH PAPER

Event detection in twitter by deep learning classification and multi label clustering virtual backbone formation

Zahra Rezaei[1] · Behnaz Eslami[2] · Mohammad Amin Amini[3] · Mohammad Eslami[4]

Received: 4 May 2021 / Revised: 15 October 2021 / Accepted: 6 December 2021
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract
The spread of social networking sites (**SNS**s) has led to the development of news and its spread. In terms of momentum produced content and topics, the gathered data from online services like twitter, are from a wide range of areas. Finding an abnormality pattern as well as content oriented planning in society, is the reason why event analysis can be of great importance. Event, in social networking sites is an interesting happening in real world which causes discussion on the topic about related issues by the users of SNSs like twitter, and it is released immediately after the discussion or with a delay. Event, changes the amount of textual data, in a way that it presents related topics in a specific time period. This event is identified by time and topic, and is related with entities like people and places. The computations related to event detection in real time is a big problem in this context. A model based on deep learning is presented in this paper. Firstly, according to the labeled data, classes are formed through classification. Later, in a flow manner, unlabeled data are presented to the model. The unlabeled data are divided into the present classes according to the model which they have been trained. If the data are higher than an identified threshold, they are assigned to a new class, and when the data are lower than the threshold, they are categorized as temporary event. Innovation of the proposed method is in two issues. First, the data in this model are semi-supervised; therefore, the labeled data are used in the first phase and the rest of the data are used in the second phase. In HAN classification phase, a module titled bag of sentence was produced for exact classification of sentences, and in the second phase the abstract concept of Virtual backbone was used to enhance precision of multi label clustering. Adding these two sections using the proposed method enhanced the precision of classification and purification of the data in unlabeled data.

Keywords Deep learning · Classification · Big data · Topic modeling · Virtual backbone · Multi label clustering

✉ Zahra Rezaei
z.rezaei2010@gmail.com

Behnaz Eslami
behnazeslami30@gmail.com

Mohammad Amin Amini
aminix@outlook.com

Mohammad Eslami
eslami.mohammad71@yahoo.com

1 Department of Statistics and Information Technology, Institute of Judiciary, Tehran, Iran

2 Department of Computer Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran

3 Department of Computer Engineering, Islamic Azad Islamic University of Jasb, Markazi, Iran

4 Department of Computer Engineering, Islamic Azad Islamic University of Qazvin, Qazvin, Iran

1 Introduction

Social networks are the media for collecting and disseminating data around the world. In times of crisis, such as floods, earthquakes, fires, and the spread of the Corona pandemic virus, these media are a valuable source for alerting and educating the public. The large statistical community of Twitter users makes it possible to monitor events significantly. Immediate detection of events is very important to react quickly to an event.

In the proposed method, two separate phases have been used in order to detect the incident quickly and with high accuracy. Classification and clustering phase in the second phase, the concept of virtual backbone and dynamic clusters is used to allocate streaming tweets. The classification, phase helps to keep the structure of tweets. The data stream from social media usually has the following characteristics: In using comments usually unsocial language is used and

Published online: 22 March 2022

there is a 140- character limitation for each comment, which results in summarizing the content of the message and sometimes blocking message delivery, a case in the point is using pls instead of, please.

Furthermore, the Twitter data are vast, and sometimes the associated documents with an event cannot be extracted easily. Some points should be taken into consideration regarding tweets: First, the information posted in social media, specifically the information with the argumentative form like weblogs has rich data.

The many challenges are in analysis in such social media such like the information is along with noise, unusual spelling and misspelling, semantic ambiguity, acronyms, a non-standard abbreviations make words and expression not understandable, thus, before the proposed phases, preprocessing phases according to the language structure needs to be done. Due to the fact that the labeled data are small and do not show the overall structure of the event pattern, semi-supervised methods should be used.

In supervised models, the number of labeled data must be sufficient to represent the data structure. Although, the number of labeled data in the realistic application is low and reaching the labels is expensive. Thus, the way to use unlabeled valuable data to train for clustering should be known.

Some methods can be used to solve the problem of label shortage and learning through a low number of labels. Examples of such methods include self-training, co-training, and generative models. The self-training method is a common method in semi-supervised learning. In this method, the data which are labeled are used for classification, and then the classifier is used for grouping of unlabeled data. Then, unlabeled data which are grouped with high certainty, are added to labeled data.

The process of adding unlabeled data to labeled ones continues so that the two groups of data are converged. If the labeled data fail to present the general structure of the data, the data face problems, as the main training of the classifier with labeled data yields weak results for unlabeled data.

In recent years, an improved version of the self-training technique called help training has been used. It is based on the idea to use a generative model to choose a part of unlabeled data to help for the training process of the main classifier.

The results reveal that this method in comparison with the self-training method yields more desirable results. Because of the generative model which uses labeled data to choose unlabeled subsets, this method has not solved the basic problem of the self-training method. Therefore, the help training model is not very useful when the labeled data fail to present the whole amount of real data. To discuss the existing events, the semi-supervised model twitter is also used.

This paper aims to provide data coverage in different classes to detect the event. In the second section of this paper, some previous studies on detecting the event are given. The proposed model and its architecture are discussed in the third and fourth sections. The results of the implementation of the model and the results of the data analysis are presented in the last two sections, respectively.

2 Related work

According to the importance of automatic event detection in Twitter, a wide variety of methods have been suggested through previous researches in relation to various techniques and characteristics including temporal approaches, topic modeling, incremental clustering, graph theory, rule mining and bursty event detection in this section which related literature in this field are presented.

2.1 Temporal based approaches

Gaglio et al. [1], employed the Soft Frequent Pattern Mining (**SFPM**) algorithm. In addition, Petkos et al. [2], employed term interestingness based methods to identify related topics in an extensive macro event from the Twitter data flow. A dynamic temporal window size was employed by twitter monitor mathioudakis and koudas [3], Twitter Live Detection Framework (**TLDF**) to detect events about their co-occurrences in real time. The TLDF can set its event detection performance based on the real amount of tweets about an event using the dynamic temporal window.

Top *k* terms from the tweets in the current time window are extracted by the term selection method in the modified SFPM in order to decrease the number of terms which should be studied. Two factors are considered to weight the values. The first factor is the term which refers to entitled entities like persons and locations by the NER Boom et al. [4], and its tf-idf score; another factor is the maximum amount of the similarity of appearance for a term in the current time window and the reference volume of the collected tweets.

2.2 Topic modeling

The philosophy of the topic modeling based approaches is that the tweets can hide some topics. Tweets are shown as a combination of topics.

The most common type of probabilistic topic model has been Latent Dirichlet Allocation (**LDA**) by Blei et al. [5], in which the topic distribution has a Dirichlet prior. The LDA is a hierarchical representation based on a three-level Bayesian model, In this way, a finite document is modeled into a set of main topics. This method is modeled on the basis of topic probabilities and generates an infinite combination of topics. Deriving good topics from the restricted context is a problem that should be solved because of the limit on the length of a tweet.

Furthermore, according to the topic modeling based approaches, too much computational price is considered as an important issue in a streaming setting, and are not effective in dealing with the events that are described in parallel Aiello et al. [6]. Stilo and Velardi [7], believe that the LDA-based methods can only be applicable offline, as the temporal dimension of the events are not taken into account in clustering approaches.

2.3 Incremental clustering

Hasan et al. [8], in their Twitter News+, used an incremental clustering based approach which solved the problem of event detection from the Twitter data flow with a less expensive solution to detect both major and minor events from the twitter data flow. Twitter News+is an event detection system which includes specialized inverted guidelines and an incremental clustering approach to detect both major and minor events in real-time with a lower cost, the most expensive operations in the Search Module and the Event-Cluster Module algorithms of Twitter News+are complexity in terms of computational of **O(1)** with parallel processing. Different filters which are employed after the production of candidate events, make a computational cost of $\boldsymbol{O(n^2)}$, in which n is the number of tweets in an event cluster. [9], presented an effective method to choose the keywords which are frequently used in twitter and are mainly linked with events which are interesting for such protests. The amount of these keywords is followed in reality to detect the interesting events in a binary classification scheme.

The purpose of the proposed method is to binarize daily count vectors of for each word pair by using a spike detection temporal filter, then this model uses the Jaccard metric to identify the similarity of the binary vector for each word pair, and the binary vector describes event occurrence. The top n word pairs are used to classify any day to be an event or non event day. The mutually generative Latent Dirichlet Allocation model (**MGE-LDA**) was proposed by Xing et al. This model uses hashtags and topics, since they both are produced mutually by each other in tweets.

2.4 Graph theory

Taking after the fruitful application of chart hypothesis in social media analysis like tweeter, there has been an inclination to utilize graph-based approaches for detecting events.

Moreover, later investigations recommended utilizing Structural Clustering Algorithm for Networks (**SCAN**) to extract the communities within the charts, social networks posts, and content as graph nodes [10].

Recent researches have recommended a named entity-based strategy considering the high computational complexity associated with generating graphs [11] After recognizing the entities, only the setting around them was considered to extract hubs, edges, and weights. Indeed in spite of the fact that keyword-based strategies speed up the graph preparing, they are less expandable due to the utilization of language or domain-specific for catchphrase extraction.

In [11], they offered a novel method that utilizes NE mentions in tweets and their entity context to create an event graph in the form of temporal. Afterward, utilizing basic graph theory strategies and a PageRank-like algorithm, they create the event graphs to identify clusters of tweets describing the same events.

2.5 Rule mining

Methods have been used in the detection of textual events, which is a method to discover the Association Rule in the temporal analysis of textual concepts [12, 13]. In [14] High Utility Pattern Mining (**HUPM**) was proposed to detect the frequent but also the high in utility item sets helped utility of terms was defined based on the growth rate in frequency. This method is based on identify the topics or events without recognizing temporal event occurrence details. Nguyen et al. [15] used a rule-based method to extract known entities for event detection.

They were represented by documents and clusters as entity-document and entity-cluster inverted indices which were used for head cluster generation. Also they are from semantical groups for used the recognition of important terms related to events so this method did not consider the relationships between words.

2.6 Bursty event detection

In data stream is able to discover not only popular topics, but also Identify bursts events generated by a large number of unusual messages in the same document in a short period of time In this article, a group of tweets are clustered in the center of the cluster. Textual and temporal features are preserved.

In simple terms, the tweets are summarized, thus reducing the size of the clusters. This method, in addition to clustering important tweet messages, maintains valuable nodes that are generated during the event. The speed and accuracy of detecting repetitive and sudden events in this method is high [11].

Real time event detection in data stream is analyzed in this paper.

Different data is generated across the Twitter social network, and feature extraction is difficult to find a definite pattern due to the large number of users. In this method, after text normalization and weighting based on systematic words, The following steps are performed in order: data stream clustering is done and effective features are extracted from the clusters, decomposition rate of a cluster and optimized of events stream with **threshold T** [16].

In this process, the relationship between topics and hashtags in tweets are modelled and it uses them as features for event discovery [17]. In another study, Azzam et al. employed deep learning and cosine similarity to find out short text posts in communities of question answering [18, 19].

Recent studies [20–22] proposed utilizing incremental clustering [23] to unravel the event evolution issue. Models are incrementally updated as unused information arrives on a stream. Such techniques may not be feasible to utilize for the Twitter Firehose because of the scale of updates.

In this paper, we unravel this issue by some deep classification modules before the virtual backbone arrangement.

This sort of linking was proposed in [24]; nonetheless, they were not able to illustrate their end-to-end approach performing in an online setting. Our strategy achieves event detection with advancement following in real-time through modeling events as the virtual spine and addressing scaling concerns with new plan choices.

3 Proposed methods

Social media are important in that each user is a potential writer. The language of the writings is closer to reality more than any other linguistic rule. This specific type of data mixture enhances event detection using models which have been made from big data analysis.

To discuss the existing events, semi-supervised model twitter is also used.

One way to figure out the reaction of the users to what is happening, is the capability to check the tweets that discuss a specific event. Therefore, in this paper, a semi-supervised method is presented which monitors related tweets about a special event, in order to make a timeline.

The main idea of the present paper is to use labeled data in the learning process of classifier. In this method, a model has been proposed for semi-supervised modeling in which classification and clustering have been combined. Semi-supervised method is combined with self training process for better understanding of classifier.

The positive point of this method is that labeled and unlabeled data are used to show the main structure of data atmosphere using clustering to make up the limitation of labeled data.

Semi-supervised clustering was employed through labeled data for instructing the clustering process. The new unlabeled data are used both for updating the classifier, and for better training of clustering methods. The overall architecture of the offline and online phases is illustrated in Fig. 1 which the phases of the proposed method are as follow:

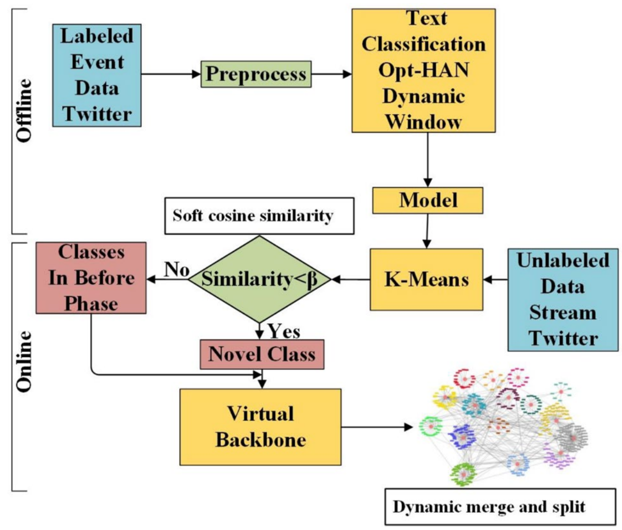

Fig. 1 Architecture of the proposed model

1. ***Data Collection*** Refers to the details of how to use datasets in the offline phase for classification and use of data stream in the online phase.
2. ***Classification*** Refers to the details of implementation method in order to text classification in three levels of word, sentence, and phrase.
3. ***Topic Modeling*** In this phase, data stream is entered into the architecture and the main topics are extracted for clustering.
4. ***Clustering*** In this phase, virtual backbone is formed and event graph is produced.
5. ***Event Detection*** In this phase, the dynamic clusters are created in each time slice and based on the overlapping events turn into sub-events, small and large clusters. Meanwhile, some tweets become very bold and pale edges. In fact, the edges describe the power.

3.1 Data collection phase

Table 1, illustrates input datasets.[1] A collection containing data for 30 different Twitter datasets associated with real world events were used in this paper [25]. The datasets were gathered within the time period of 2012 and 2016, using the streaming API with a set of keywords that are listed below.

[1] https://figshare.com/articles/Twitter_event_datasets_2012-2016_/5100460

Table 1 Tweets distribution on 2012–2016

Title	From	To
Euro2012	07/02/2012 11:33	07/02/2012 13:12
Obama-Romney	11/05/2012 05:34	11/05/2012 09:36
Mexican-Election	07/01/2012 18:43	07/01/2012 20:24
Boston Marathon Bombing	04/16/2013 01:10	04/16/2013 01:28
Ebola	07/01/2014 12:00	07/12/2014 07:59
Hong Kong Protests	09/26/2014 12:45	10/14/2014 10:51
Ottawa Shooting	22/10/2014 03:32	22/10/2014 17:56
Nepal Earthquake	04/25/2015 06:18	05/10/2015 01:24
Hurricane Patricia	10/24/2015 12:04	10/26/2015 12:57
Charlie Hebdo	01/07/2015 10:49	01/07/2015 02:27
German Wings-Crash	03/24/2015 10:34	03/24/2015 01:03
Brexit	02/24/2016 12:54	04/13/2016 06:43
Brussel Airport Explosion	03/22/2016 08:28	03/22/2016 13:54
Hijacked Plane Cyprus	03/29/2016 07:16	03/29/2016 10:38
Irish-ge16	02/03/2016 11:05	02/15/2016 23:56

From the twitter datasets 150000 comments were extracted **randomly** from twitter event dataset comments. Furthermore, such datasets were employed to compare comments with Twitter with various points of views.

Furthermore, these datasets are used to evaluate the method for each class of 1000 tweets which were in English.

From among 150000 tweets, 135000 were chosen for the training phase and 15000 were chosen for the test phase in a **tenfold** manner. These analyses are designed for major events in these years.

It seems necessary that, less important events may have occurred in these years due to their coincidence with the stronger event becoming less faded and not included in the standard dataset, but added phase down of the virtual backbone of those events has been shown, such as the 2013 Boston Marathon bombing that occurred in China earthquake on the same date, and tweets were released but not collected due to the importance of the data bombing event, but in the backbone Virtual is defined as a smaller cluster.

3.2 Classification phase

Since in the dataset of Table 1, there are labeled data, we used the Hierarchical Attention Network (**HAN**) method to categorize the data. Indeed, HNN was the method used for data classification, when word2vec is made; this file is used for further steps of study.

HAN has two main features: first, a hierarchical structure that illustrates the hierarchical structure of documents; and second, two steps of attention mechanisms which are employed at the word and sentence level, and which enables the network to perform differently to more and less important content when constructing the document representation.

In addition, the HAN network contains a few parts including; a word sequence encoder, a word level attention layer, a sentence encoder and a sentence level attention layer.

Based on HAN, considering sentence and documents structure in modeling, results in better representation of document structure in the architecture of the model which it's called Opt-HAN (Table 2). The attention mechanism in NLP gives the capability to the model to learn what to pay attention to according to the input text. In fact, it does not present the full source text into a fixed length vector like standard RNN and LSTM.

The difference of HAN and Opt-HAN in this paper, employs context in order to detect when a specific sequence of tokens is related rather than easily filter different mixtures if signs in a text. Our method has enhanced the previous methods. All words do not cooperate in presentation of a sentence similarly.

Therefore, via attention mechanism, we identify the words which are important in understanding of the text, and then these meaning rich words are juxtaposed in the form of a sentence vector.To detect sentences which work as clues for classification of a text, the importance of each sentence is determined via a vector and using a context vectors.

Then, if **n** sentences are considered in **m** documents, a greater possibility of multi label classification of the text can be obtained, that is how much of a text presents a special topic and how much of that text shows another topic.

It is important to note the level of dependence of text on each class.

The length of the document affects the number of attention modules in each document.For instance, in short text with only one sentence, this section is deactivated. Also in long text Attention module of document level is used.

So, the attention module of sentence level between two classes reaches lower levels than threshold. Thus, depending on the size of $(\frac{n}{2} + 1)$ of the previous and next sentences, sentences with low dependency are investigated using RNN short term memory.

We optimized HAN, in which we benefited from novelty in technique in order to increase accuracy and finding meaningful correlation in text. In this optimization, the new modules of sentence segmentation or section to find cohesion in meaning was added to previous or first modules, i.e. word encoder, word attention, sentence encoder and sentence attention.

In other word the new or second module is segment sentence encoder and segment sentence attention namely bag of sentence modules that illustrated in Fig. 2.

The performance of this added phase is to segment meaningful sentences, this would increase the percentage of true allocation to specific class.

Since the mentioned sentences by publics in social media such as twitter may not have accurate base, the whole written

Table 2 Details of the proposed method

Assumption GRU:
GRU gating modules (2 sub modules) //tracking the state of the sequences without memory cell
Word2vec row document converts to a vector representation
Control how much information is updated at time t
r_t: reset sub module //how much the past state contributes to candidate state
h_{it}: hidden representation
u_{it}: importance of the words as similarity
u_w: word level context vector //randomly initialized and learned during process
u_s: sentence level context vector
α_{it}: normalization importance weight through a softmax function
z_t: update sub module //how much past information is kept or is added
$z_t = \sigma\left(W_z x_t + U_z h_{t-1} + b_z\right)$
h_t: new state // $\widetilde{h}_t = \left(1 - z_t\right) \odot h_{t-1} + z_t \odot \widetilde{h}_t$// h_{t-1} previous state
$\widetilde{h}_t$: current new state //new sequences information, candidate state, computed
in a mechanism similar to RNN
$h_t = \tanh\left(W_h x_t + r_t \odot \left(U_h h_{t-1}\right) + b_n\right)$
w_t: window sequence //dynamic candidate state, forward and backward parallel
//weight of different entities (words or sentences)
x_t: sequence vector at time t
S_i: sentence vectors
L: number of sentences //S_i
T: number of words t_i of S_i // $t \in [1, T]$
t_i: word in i'th sentences
if r_t=0 then
forget the previous state
$r_t = \sigma\left(w_r x_t + U_r h_{t-1} + b_r\right)$

text probably would not be correlated to different concepts rather than one specific concepts.

Detection of relationship between sentences requires finding some segment to consider the allocation of sentence to the specific class, which is done by the 3rd module of sentence segmentation. It is good to be mentioned that the percentage of allocation is dynamic.The structure in HAN is hierarchical, and its mechanism at first is at word level and then at sentence level and finally it is bag of sentences.

In fact, this is the reason why the importance of words and sentences is dependent on the field in which they are used, so that the same word or sentence may have a completely different value in a different field.

Therefore, three levels of mechanism are taken into account, and this gives us the possibility to differentiate more important parts of the text from less important parts while making a specific presentation of a text.

The general structure of the HAN is presented in the Fig. 2, which includes various components: An Encoder and Attention at word level and an Encoder and Attention at sentence level and finally adding Encoder and Attention to the whole document.

Through hierarchical usage of word vectors (Word2vec) based on the performance of the mechanism Encoder and Attention at word and sentence levels reciprocally (once from the end to the beginning and once from the beginning to the end), the similarities are identified, and in each time of learning its precision in detection is increased.

In general, it can be said that teaching of HAN network, is highly dependent upon word vectors (Word2Vec), as different embeddings are tested the resulty was wonderful that a greater number of dimensions of the embeddings are not always a guarantee for better results [26]. Actually, In this paper based on the Table 3, the proper embedding is 100.

In fact, what is important in this proposed method, is substitution of its default function from tanh to sigmoid which affects the precision or error of the output [27]. Details of the proposed method are presented in the following section. Additionally, the most important HAN's method parameters are listed in Table 3.

The change which has occurred in the HAN model is explained here. First, in the section which deals with attention on word in the sentence, mutual attention is conducted simultaneously.

In addition, dynamicity window is added to the Encoder module, which is based on summarizing topic (Table 4).

All words are not in the same level of importance; therefore, the words that have more important roles have

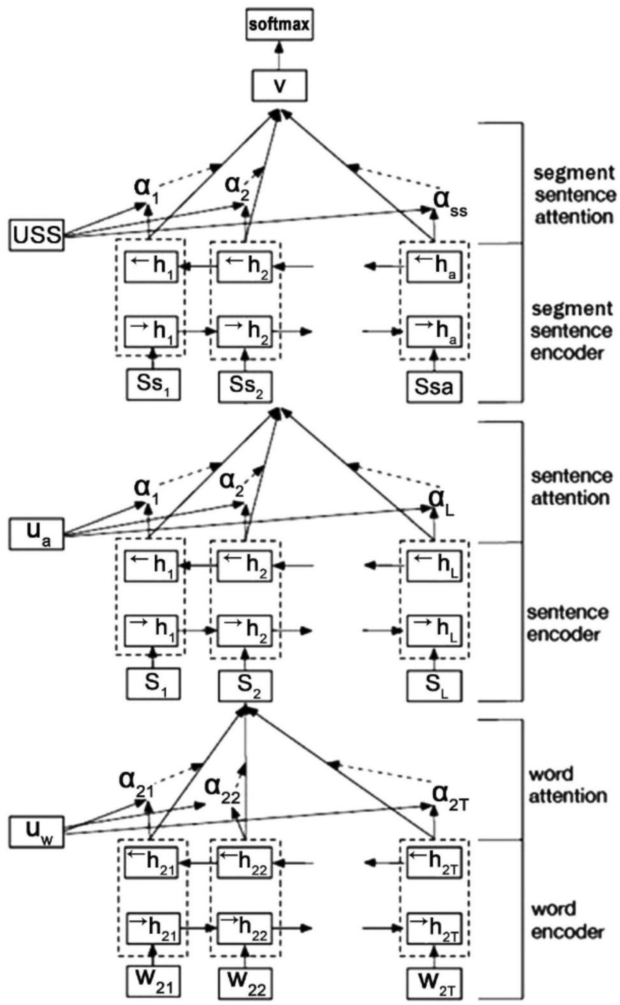

Fig. 2 The architecture of changed Opt-HAN method

Table 3 HAN hyper parameters

Optimization	Loss Function	Learning Rate	Embedding_dim	Pad_seq_Len
Adam	Sigmoid	0.5	100	150

a higher weight in classification, that is the issue is attention weight modules. Also, All words are not identically important in realizing the document; therefore, the sentences which have more important role, should be classified with a higher weight, that is the issue of attention weight modules is at work again (Table 5).

In the sentence section, the dynamic window moves between the previous and next sentences identically, and chooses the sentence which conveys the highest level of attention. The *t_index* shows the dynamicity of the window while investigating the importance of the sentence (Table 6).

Table 4 Word Encoder modules

w_{it}: convert word to vector through an embedding matrix //W_e
$x_{ij} = W_e w_{ij}$// bidirectional GRU concurrent Elapse time to get annotation of words by topic attention, summarization information, from both direction for word incorporate the contextual information in the annotation
$\overleftarrow{f}$: forward GRU //reads the sequences *si* from w_{i1} to w_{it}
$\vec{f}$: backward GRU //reads the sequences *si* from w_{it} to w_{i1}
For i = 1 to s_i, $s_i \in [1,\text{point}]$ //concurrent with dynamic sequence window
$x_{it} = W_e w_{it}, t \in [1,t]$
$\overleftarrow{h_{it}} = \overleftarrow{GRU}(x_{it})$
$\overrightarrow{h_{it}} = \overrightarrow{GRU}(x_{it})$
$h_{it} = \widetilde{w}_t\left(\overrightarrow{h_{it}}, \overleftarrow{h_{it}}\right)$// based on topic attention

Table 5 Word and Sentence Attention modules

$u_{it} = \frac{1}{1+e^{-(W_{w,s}h_{it}+b_{w,s})}}$
$\alpha_{it} = \frac{exp(u_{it}^T u_{w,s})}{\sum exp(U_{it}^T u_{w,s})}$
$s_i = \sum \alpha_{it} h_{it}$

Table 6 Sentences Encoder modules

For all S_i
$\overleftarrow{h_{it}} = \overleftarrow{GRU}(S_{it}), i \in [1,L]$
$\overrightarrow{h_{it}} = \overrightarrow{GRU}(S_{it}), i \in [1,L]$
$h_{it} = \widetilde{w}_t\left(\overrightarrow{h_{it}}, \overleftarrow{h_{it}}\right)$// based on topic attention in sentences

Table 7 Phrase Encoder modules

For all S_i
$\overleftarrow{h_{it}} = \overleftarrow{GRU}(S_{si}), i \in [1,L]$
$\overrightarrow{h_{it}} = \overrightarrow{GRU}(S_{si}), i \in [1,L]$
$h_{it} = \widetilde{w}_t\left(\overrightarrow{h_{it}}, \overleftarrow{h_{it}}\right)$// based on topic attention in segmentation of sentence

Assuming the vectors of the S_{si} statement, the same vector can be obtained for the document. Using a bidirectional GRU, the encryption of the phrase was done as follows: By connecting $\overleftarrow{h_i}$ and $\overrightarrow{h_j}$. As hi = [$\overleftarrow{h_i}$, $\overrightarrow{h_j}$] the interpretation of the expression was obtained. h_i, in short, contains adjacent phrases around the word i, although the focus will still be on h_i according to Table 7.

After the above mentioned stages, important sentences are extracted based on the importance of words and sentences. Now, the sentences which are more important are

Table 8 Phrase Attention modules

$u_{it} = \frac{1}{1+e^{-(W_{ss}h_{it}+b_s)}}$
$\alpha_{it} = \frac{exp(u_{it}^T u_{ss})}{\sum exp(U_{it}^T u_{ss})}$
$\Theta_i = \sum \alpha_{it} h_{it}$

arranged based on the topic of summarizing in the document and based on the weight. Then, the general tendency of the document is identified according to the previous four steps as mentioned in Table 7 phrase encoder assuming (S_i) expression vectors can be similarly accessed by a vector for a document. Using a bidirectional GRU, the phrase was encoded as follows: Document attention module, to determine documents that are involved in text classification, using the attention mechanism again by introducing a context vector (u_s), through a vector the significance of each document is determined by the relation the following leads. The important note in the phrase segmentation is to determine the degree of importance of the phrase in the document. Not all sentences are relevant to the meaning of the document in a sense, so the sentences that have a more important role to play in the higher weighting category, namely the discussion of the attention weighting module, that mentioned in Table 7.

Actually the important terms were extracted after the lifting steps based on the meaning of the word and sentence. The sentences that matter most are sorted by weight according to the topic of the summary in the document, and the overall trend of the document is determined by the three preceding sections, respectively (Table 8).

3.3 Dynamic topic modeling phase

Non-negative Matrix Factorization (**NMF**) Lee and Seung [28] which is a linear algebraic model introduces high dimensional vectors into a low dimensional image. Similar to Principal Component Analysis (**PCA**) Wold et al. [23], NMF takes the fact into consideration that the vectors are non negative.

The two matrices of **W** and **H**, would be obtained through original matrix **A**, in which **A = WH**.

Also, NMF has an inborn clustering property and **W** and **H** represent the following information:

A (Document-Word Matrix) input that shows which words appear in which documents.

W (Basis Vectors) the topics (clusters) are elicited from the documents.

H (Coefficient Matrix) the membership weights for the topics in each document.

W and **H** are calculated by optimizing an objective function (like the *EM algorithm*), and updating both **W** and **H** iteratively until they are converged.

In this way, Table 9 shows the configuration of the NMF method for extracting topics by evaluating number of topics dynamically, based on the **TC-W2V** coherence measure.

3.4 Clustering and virtual backbone phase

In order to model relationships between terms in the contexts, in Clustering and virtual backbone phase, an event

Table 9 NMF Topic modeling configuration

Initializations:
Step1- Applying Term Weighting with TF-IDF // Create TF-IDF-normalized document-term matrix as a pickle (**.pkl**) file
Step2- < Finding the **best k-value** for number of topic dynamically >
Create the Topic Models by pre-specifying an initial range of "sensible" values:
decomposition NMF init = **"nndsvd"** // better for sparseness
Build a Word Embedding model from all documents in the input file
Selecting the Number of Topics by implement a simple version of the **TC-W2V coherence measure**
Number_of_Topics = Select max value as a number of topics
Step3- < Extracting Topics >
Examine the Final Model
Number of Topics: {**"Number_of_Topics"**}
Run to extracting Topics

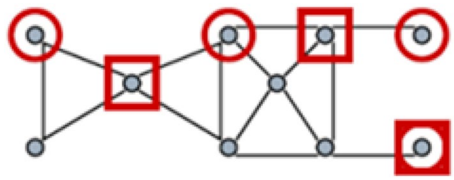

Fig. 3 Maximal Independent Set by Johnson et al. [31]

graph is produced. A graph theory which will be considered as event candidates is used to divide the graph into subgraphs. Tweets which deal with similar events usually have similar keywords.

On the opposite, tweets which are related to different events usually have different keywords. This can be seen by stronger relations between nodes which deal with the same events in the event graphs.

In other words, in cases when tweets about similar events connect the edges, their weights are higher compared with the time when edges connect terms from tweets about different events. The edges are divided to recognize such edges that, if the edges are removed, they will divide the large graph G into sub graphs.

In some cases, the achieved graphs at a definite time window are disconnected. In other words, all the nodes in the graph are not related to each other. At first, the generated graph is processed through analysis of the link between different nodes.

Accordingly, the nodes which are connected with few other nodes are deleted from the graph. At this phase, the main event graph is classified into a number of sub graphs which include keywords that are highly linked. In the present detection approach, it is considered that events from different sub graphs are not related to each other. Therefore, in the event detection sub module, each sub graph is processed distinctly from others.

If you assume that each input sample (the tweets) is entered to the model for clustering, and if it is considered that the sample is allocated to more than one class based on the classification model, then the following definitions should be presented regarding backbone: Independent Set (**IS**) by Griggs et al. [30], is taken as a subcategory of cluster sample or graphs which have no two members beside each other. According to Johnson et al. [31], an independent set with the highest number of members in a graph is known as Maximal Independent Set (**MIS**). In other words, if every cluster sample is added, the independence feature of an independent set is lost. Therefore, each sample which is not a member of MIS is next to the sample of MIS's cluster which is presented in Fig. 3.

Dominating set (**DS**) is another subset of cluster sample which is like **V**; every sample belongs to V-DS and it is neighboring to at least one of the nodes in DS. Thus, it can be claimed that MIS itself is a DS but DS is not an MIS. DS is usually employed to relate sample to the cluster heads in clustering of twitter.Finding a head set which has the least size or MDS is a NP-Hard issue. An instance of DS is Connected Dominating Set (**CDS**) which is a dominating set in which the inductive sub graphs are linked. A CDS which has the minimum possible cluster in the model is called Minimum Connected Dominating Set (**MCDS**). According to Cheng et al. [25], a virtual backbone can be considered. CDS is a set of network nodes which are linked in a way that every node belongs to CDS subset or at least is next to a node in CDS.

This structure can be employed to begin a virtual backbone to find the best cluster in model, as it is connected. Since this structure is interconnected, it can be expressed as a virtual backbone as an abstract concept of belonging to one or more clusters. For example, if there are N main topic in the input data, that is, if we have N classes of data, each input tweet will be subdivided into at least one or more classes by topic, and if no classes belong to it, that is, less It is over the threshold and is designated as the new uncertain class. In fact, after the online phase is complete, when the stream data is entered into the proposed model, clusters are formed based on k-mean topic modeling.

At this stage, each document will have an event-based topic will belong to more than one cluster which this is a high-level abstract of multi-class clustering.

The difference between the proposed method and the multi-cluster clustering is that in this method each document can belong to one or more clusters and does not have to belong to a particular cluster.An important concept of the virtual backbone is that it is important that during data entry into the model, the degree of edge may be low or high, and that sometimes the relationship of a topic to a cluster may be severed.

3.5 Event detection phase

An event is defined as an incidence which happens at a place or time. In addition, the event may occur in various places within several days like infectious diseases. A group of expressions the frequency of which is increased intensively can be used to detect an event during the analysis period, one or several times. Hashtags provide an overview on topics in tweets, a method for this paper is using a technique for outstanding topics. The first phase is classifying hashtags, so they can be used as the event index.

Based on the fact that the proposed model of classification is multi label classification, noting that each input sample belongs to more than one degree, this virtual backbone is formed using the model, and is converged to one of the clusters according to the results of belonging degree clustering. Addition of a new class is another important point.

In some cases, the current discussions cannot be subsumed under one of the present classes, and because of low level of similarity to the current clusters, a new cluster should be made. Due to the variability and the dynamic nature, the clusters of new class are made. Creating event in twitter is conceptually different from trend.

An event is usually created in twitter, and sometimes because of high frequency of retweeting, it become a trend. The topic modeling phase helps event determination in this model. Based on the fact that the frequency of topics in each cluster is different, event detection in limited range is very difficult.

In addition, the number of topics is also considered to be dynamic, and a cluster may contain two or several important topics.

When a topic changes compared with the previous topics, a significant change happens in the distribution of data in each cluster, then the cluster is defined as a new event. Minimum connected dominating set is used to minimize events over time because without the assumption that the number of nodes is minimal, due to the large variety of events occurring in this large set of nodes, it is very difficult to control the probabilities of excessive event of nodes. Actually soft cosine similarity is a measure of how similar the documents are regardless of their size and mathematically, the cosine criterion measures the angle between two vectors (the number of words of two documents) in a multidimensional space.

A cosine similarity is considered 0.5, that is, if a topic has an over-threshold similarity, it belongs to the cluster, otherwise it will be compared to another cluster.

Tweets may belong to several clusters which the similarity of each tweet is measured with the cluster head, and at each stage within each cluster is determined by the importance of the intra-cluster dominant topic. For example, in the Obama presidency, as president, he is the cluster head in this cluster of important election issues, the Democratic Party, the Republican Party, Freedom, New tweets on election issues will be compared with the head in the cluster.

Another point is the process of merging and splitting the clusters formed in the minimal connectivity set after each time period with respect to the data extracted, if the number of clusters exceeds the threshold, similar clusters whose cosine similarity criterion is greater than 0.7 are merged.

For example, in 2013, hurricane, flood, and earthquake events each formed separate clusters. After three months of data entry, all of these events can be merged as a natural disaster cluster. It did the separation threshold, but because it aims to determine a new class, the separation is done by the same criterion.

Table 10 Result of classification before bag of sentence in Opt-HAN method (Fig. 2)

Metric	Section 1	Section 2
Batch Size	64	128
Learning Rate	0.001	0.001
Accuracy	0.958	0.967
Kappa	0.903	0.928
Recall	0.919	0.967
Precision	0.913	0.965
F1-Score	0.919	0.967

3.6 Computational complexity

3.6.1 Offline phase

Twitter is considered an SMS text-based service with some restrictions in terms of tweet length. The original Tweet's length was 140 characters which over time evolved into a maximum amount of length of 280 characters and in this paper, the average number of tweets in the classification phase was 160 characters.

As a standard page (**SP**) has a strictly defined length of text, therefore one standard page is comprised of 1500 characters with spaces. The total amount of time which has been spent on classifying, according to the standard page lenght which has 1500 words, while the whole test document had 2056 pages, our model managed to analyze it in 20.148 s. More specifically, each page has taken 0.014177043 s to analyze. On average, it takes 0.00150 s to classify each tweet note which this time is related to the offline phase. In terms of computational complexity, deep learning-based methods for classifying short documents have been relatively successful.

3.6.2 Online phase

Topics in the previous step are extracted very quickly with the topic modeling algorithm, entered the virtual spine, and assigned to nodes based on the classification which is in the predicate phase.

So it is noticeable the advantages of the suggested algorithm are related to limited computational complexity and the number of topics that do not have to be defined in the online phase since 87% of tweets is placed in the virtual backbone based on the previous classification and assigned immediately. All in all the timing is evaluated real-time.

Table 11 Result of classification phase

Metric	Section 1	Section 2
Batch Size	64	128
Learning Rate	0.001	0.001
Accuracy	0.9823	0.9856
Kappa	0.9841	0.9864
Recall	0.9819	0.9856
Precision	0.9825	0.9786
F1-Score	0.9823	0.9856

Table 12 Classification methods AUC comparision

Method	AUC
NB	0.894
SVM	0.678
MLP	0.923
LDA	0.919
RT	0.923
DT	0.900
KNN	0.690
HAN	0.942
Opt-HAN	0.985

4 Result

The output of the two-phase results is explained separately. In the first phase, the classification result before the HAN method is improved and after the addition of the third layer is explained in Tables 10 and 11, and 12 shows a comparison with other methods performed on this database that shows that the method improvement has made a proposal. In the second phase, the clustering results and the display of the new class are presented.

Actually, Table 11, shows the result of proposed method in the classification phase.

As it can be seen in the batch size **128** and the learning rate **0.001** the best output was obtained. The classification and modeling phase is before virtual backbone phase means that the output of this stage will be the main nodes in the virtual spine. In the online phase, unlabeled data will be placed on the current nodes upon entry into the model upon entry into the model.

As explained in the before tables, each word pair of an integer vector is converted to a vector with binary values, which is defined as Bag Of Words (**BOWs**).

BOWs will be used to calculate the binary vector jacquard similarity index with the binary vector score of the events according to the number of word-pair spikes corresponding to the event days.

Cosine similarity, Mutual Information, distance correlation, and Pearson correlation can also be used to select features in a variety of classification methods and report the best output of each method in the table below is Jaccard Similarity in future selection that was the better than others [32–37].

In the [38], the focus has been on the virtual backbone, and the changes in the classification phase compared to the Han model have only been the addition of the third phase in the document, but with the addition of implementation details in the present paper and considering the dynamic window phase, It has also been found that the addition of the merge and split phase and the combination of the virtual backbone is more dynamic and more detailed events can be detected in a shorter period of time.

In this way according to the Table 12, the proposed method in the supervised learning phase has shown better performance than the supervised learning methods due to the better focus of the proposed method on vocabulary, sentence and phrase and stronger extraction of local features in learning in It is a modeling process.

A comparison of classification AUCs using word-pairs extracted by different feature selection methods and proposed method (Opt-HAN).

Sink node which changes dynamically was introduced in this paper. In fact, the sink node has different edges as hot topics which show events and finally is assigned in each window with a high possibility to one or more events.

In this paper, the correlation matrix was used to calculate the relationship between the topics. The number of topics per cluster is different and the number of small events that occurred over a given time period was small compared to the major issues, so small clusters will be formed alongside the large clusters.

Considering the change of each topic in comparison with the previous topics, if a significant change happens in the distribution of data in each cluster, then the cluster is regarded as a new event.

As shown in Fig. 4, the number of tweets were identified as novel classes, which, although related to other classes but their correlation in previous classes are low and were identified as novel classes.

In this step, the novel class can be classified as one or more new classes based on the topic. The other issue is that the data of 2012–2016 datase which mentioned as a footnote before, is automatically entered in the model based on offline, but at the same time, the current online stream of API data that the novel class data is entered to the model and is changed probabilities of tweets.

The novel classes are related on events that isn't important in timespan or they are less important than previous event classes. Figures 4 and 5, present the number of

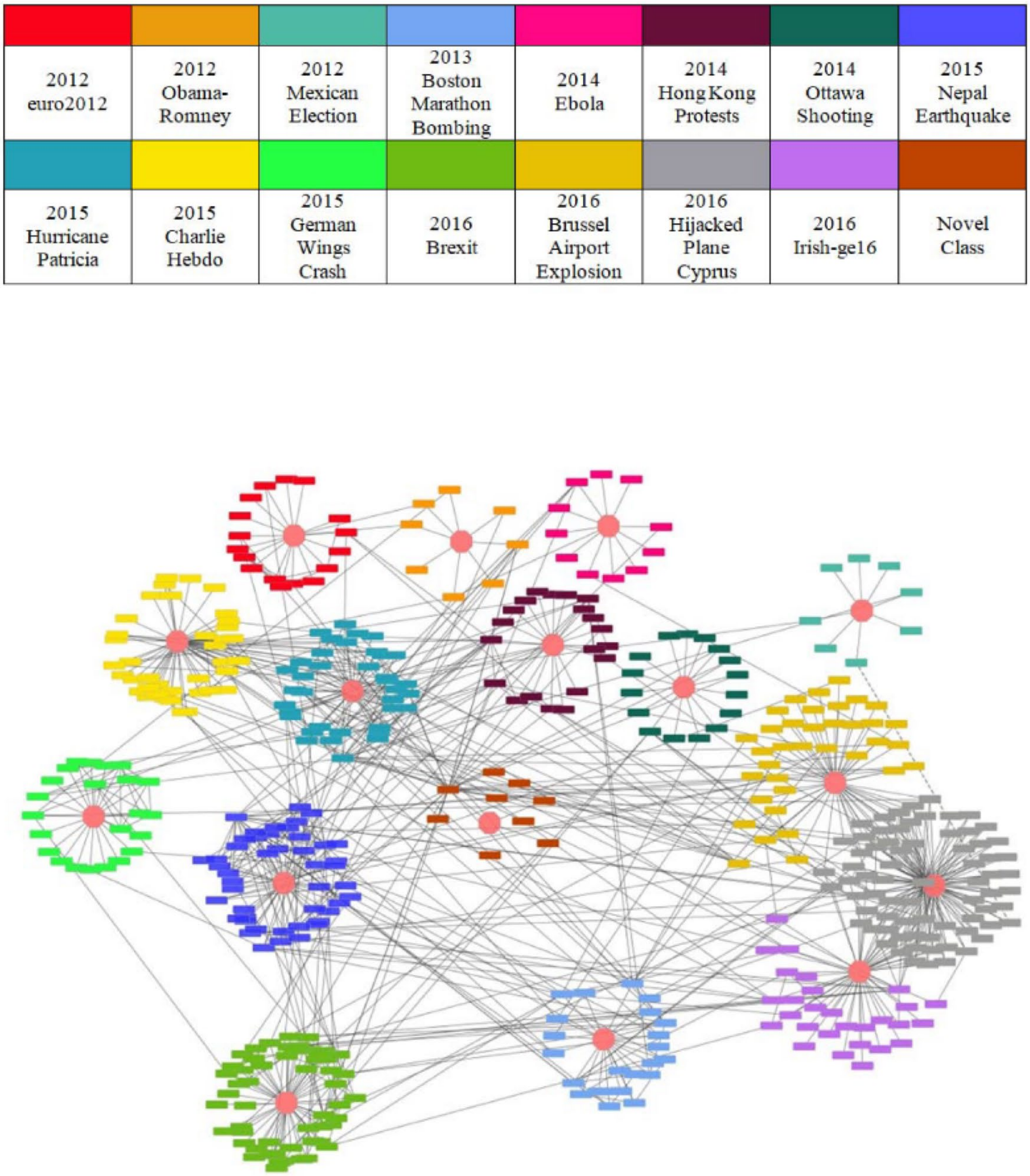

Fig. 4 Connections of events in Virtual backbone in two phase (offline in dataset, and data stream api twitter in timespan 2012–1016)

effective topics in each event in the stream data and graph of effective topics.

Actually, Effect virtual backbone in this model is returned to online streaming phase that is changed correlation of tweets in previous classes and is shown there are others event as novel class that in Fig. 4, is shown brown color.

As a new topic model, TopicVB is presented so that the Correlated Topic Modeling is available for phrase level topics.

More contextual information of phrases is employed by TopicVB; then, the phrases are put in dynamic virtual backbone; thus, at phrase level TopicVB can present at phrase level. It was shown that the correlated topic modeling on phrases is more fruitful for interpretation of the main themes of a corpus.

To be used for large datasets, the efficiency of TopicVB is optimized in future.

If the prepared data as described in subsection are presented, then in the first stage, better correlated phrase level topics $\beta\rho$ will be instructed.

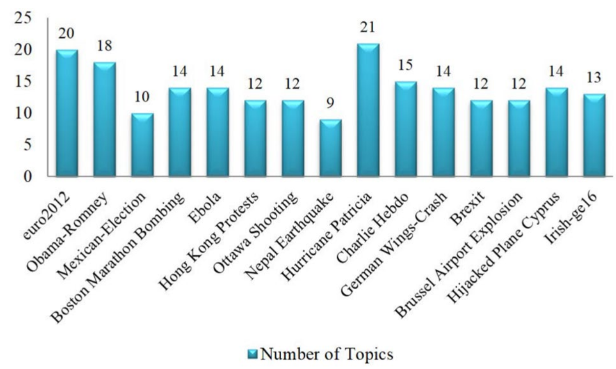

Fig. 5 Dynamic value of topic's number

The contextual information of phrases is formed by words in the same document and their component words within semantically coherent links.

The component words in semantically coherent links were modeled in the previous subsection. As with words, the phrases and words in a same document d have the same topic parameter η d, which is a K-dimension vector taken from a Gaussian Distribution N (μ, Σ). Σ, which presents the correlation between topics is the covariance matrix.

If we want to obtain the correlation for$\beta\rho$, we should consider that it cannot be obtained fromΣ, which was obtained in the first stage, this is because of the fact that Σ includes the impact from word topics. So, with $W^{(p)}$ and $\beta\rho$ in hand, the variation inference can be used obtain $\Sigma^{(p)}$. Thus, the correlation matrix is calculated through the following equation $corr_{ij}^{p} = \frac{\Sigma_{ij}(p)}{\sqrt{\Sigma_{ii}(p)\Sigma_{jj}(p)}}$ (Table 13).

5 Conclusion

A semi-supervised framework is proposed in the present paper for detection of data events of twitter. The classification model of HAN is used for categorization. To make the model 150000 tweets are introduced to the model and then, stream data are introduced without any labels to the classification phase at the same time span, and they make a new group or are subsumed under previous clusters. At this stage, virtual backbone was used for the classification phase, in a way that the stream data can be grouped into one or some classes with different grades. This grade of belongingness increases or reduces with time. When the virtual backbone is formed, topic modelling with dynamic k with stream data is utilized. In fact, the events of each class which have been more commented are detected as the main topic. In this phase, topic relations between topics are recognized for detection of main events. Finally, based on the time span, the event clusters are formed and important topics are also investigated.

All in all, the proposed method shows the improvement of clustering compared to the previous paper because of these reasons: increasing the accuracy of the classification phase due to the dynamic window at all 3 phases of the classification, defining the standard of soft cosine similarity writing similarity in K-means, adding the third phase and adding the combination phase and dynamic merge and split. So the flexibility of this method increases in the sequence of events in this highly improved method, and it shows a better display of events.

Table 13 Event Detection by Virtual backbone and Deep learning

Door list is hot topic that event class label
Dee list is relation hot topic
Algorithm: **EDVB DEEP**
Three base in EDVB (Classification Opt-HAN, Topic Modeling and Kmeans Clustering)
1: Input: Stochastic Graph $G < V,E >$, Tweet Stream with name "t", Insert
real time
to graph,V is event classes and create by hot topic of Twitter in t time slice
2: Output: Connected Dominating set (called backbone formation) multi
label classification
Begin Algorithm
Step1- <construct virtual backbone phase in stream tweets> after
classification of offline mode
Let SN be Sink Node;
Let VS be the class node of multi label classification for any event node;
Start from VS to create a backbone for multi label classification; Vi be
selected node *ith*;
Door_list_labeled = Door_list_determined = VS;
Dee_list = Neighbors (VS);
Door_list_labeled = Door_list_labeled ∪ Vi;
Dee_list = Dee_list ∪ Neighbors (Vi);
Step2- <Learning phase>
begin of while
$i = 1$;
While (Sink node did not Dominator yet and $i \leq 3$ and month(3×30) and $cosine \geq 0.5$) do:
Select one if the elements Door_list_labeled;
Calculate cosine similarity (mapping idea)
$$(a \cdot b) = \frac{\sum_1^n a_i b_i}{\sqrt{\sum_i^n a_i^2}\sqrt{\sum_i^n b_i^2}}$$
If (Ci > Cavg) // $corr_{ij}^{p} = \frac{\Sigma_{ij}(p)}{\sqrt{\Sigma_{ii}(p)\Sigma_{jj}(p)}}$
Create edge in topic and increase weight of edge;
else New cluster;
Door_list_ determined = Door_list_labeled ∪ Vi; // based max
probability Action Vector
Dee_list = Dee_list ∪ Neighbors (Vi);
$C = C + 1$;
E[c] = Wi;
$i = i + 1$;
end of while
Re-Clustering (Merge of clusters)
if (Vi is Dee and hasn't neighbor of Door) **then**
Vi will be one of elements Door_list_determined and isn't
correlated hot topic;
End of Algorithm

References

1. Gaglio S, Re GL, Morana M (2016) A framework for real-time Twitter data analysis. Comput Commun 7:236–242
2. Petkos G, Papadopoulos S, Aiello L, Skraba R, Kompatsiaris Y (2014) A soft frequent pattern mining approach for textual topic detection. In: Proceedings of the 4th International Conference on Web IntelligenceMining and Semantics (WIMS14), 25–40
3. Mathioudakis M, Koudas N. Twittermonitor (2010) Trend detection over the twitter stream. In: Proceedings of the 2010 ACM SIGMOD International Conference on Management of data, 1155–1158.

4. Boom CD, Canneyt SV, Dhoedt B (2015) Semantics-driven event clustering in twitter feeds. In: Making Sense Of Microposts workshop, 1395: 2–9
5. Blei DM, Ng AY, Jordan MI (2003) Latent dirichlet allocation. J Mach Learn Res 3:993–1022
6. Aiello LM, Petkos G, Martin C, Corney D, Papadopoulos S, Skraba R, Göker A, Kompatsiaris I, Jaimes A (2013) Sensing trending topics in Twitter. In: IEEE Transactions on Multimedia, 15: 1268–1282
7. Stilo G, Velardi P (2016) Efficient temporal mining of micro-blog texts and its application to event discovery. Data Min Knowl Discov 30:372–402
8. Hasan M, Orgun MA, Schwitter R (2017) A survey on real-time event detection from the twitter data stream. Journal of Information Science
9. Hossny AH, Mitchell L (2019) Event detection in Twitter: a keyword volume approach
10. Schinas M, Papadopoulos S, Petkos G, Kompatsiaris Y, Mitkas PA (2015) Multimodal graph-based event detection and summarization in social media streams. In Proceedings of the 23rd ACM international conference on Multimedia (pp. 189–192)
11. Edouard A, Cabrio E, Tonelli S, Le Thanh N (2017) Graph-based event extraction from twitter. InRANLP17-Recent advances in natural language processing
12. Adedoyin-Olowe M, Gaber MM, Stahl F (2013) Trcm: a methodology for temporal analysis of evolving concepts in twitter In International Conference on Artificial Intelligence and Soft Computing (pp 135–145) Springer, Berlin
13. Adedoyin-Olowe M, Gaber MM, Dancausa CM, Stahl F, Gomes JB (2016) A rule dynamics approach to event detection in twitter with its application to sports and politics. Expert Syst Appl 15(55):351–360
14. Choi HJ, Park CH (2019) Emerging topic detection in twitter stream based on high utility pattern mining. Expert Syst Appl 1(115):27–36
15. Nguyen S, Ngo B, Vo C, Cao T (2019) Hot topic detection on twitter data streams with incremental clustering using named entities and central centroids. In2019 IEEE-RIVF International Conference on Computing and Communication Technologies (RIVF) 2019 (pp. 1–6). IEEE
16. Comito C, Forestiero A, Pizzuti C (2019) Bursty event detection in Twitter streams. ACM Trans Knowl Discov Data (TKDD) 13(4):1–28
17. Xing C, Wang Y, Liu J, Huang Y, Ma WY (2016) Hashtag-based sub-event discovery using mutually generative lda in twitter. In: Association for the Advancement of Artificial Intelligence, 2666–2672
18. Azzam A, Tazi N, Hossny A (2017) A question routing technique using deep neural network for communities of question answering In: database systems for advanced applications. Springer International Publishing, Berlin, pp 35–49
19. Azzam A, Tazi N, Hossny A (2017) Text-based question routing for question answering communities via deep learning. In Proceedings of the Symposium on Applied Computing, SAC ’17, New York, USA, ACM, pp. 1674–1678
20. Becker H, Naaman M, Gravano L (2011) Beyond trending topics: Real-world event identification on twitter. In Proceedings of the International AAAI Conference on Web and Social Media 5, 1
21. Osborne M, Moran S, McCreadie R, Von Lunen A, Sykora M, Cano E, Ireson N, Macdonald C, Ounis I, He Y, Jackson T (2014) Real-time detection, tracking, and monitoring of automatically discovered events in social media. In Proceedings of 52nd annual meeting of the association for computational linguistics: system demonstrations (pp. 37–42)
22. Petrović S, Osborne M, Lavrenko V (2010) Streaming first story detection with application to twitter. InHuman language technologies: The 2010 annual conference of the north american chapter of the association for computational linguistics (pp. 181–189)
23. Hasan M, Orgun MA, Schwitter R (2016) TwitterNews: real time event detection from the Twitter data stream. PeerJ PrePrints 4:e2297v1
24. Lee P, Lakshmanan LV, Milios EE (2014) Incremental cluster evolution tracking from highly dynamic network data. In 2014 IEEE 30th International Conference on Data Engineering (pp. 3–14). IEEE
25. Cheng X, Huang X, Li D, Wu W, Du DZ (2003) A polynomial-time approximation scheme for the minimum-connected dominating set in ad hoc wireless networks. Netw An Int J 42:202–208
26. Söderman M (2018) A study of hierarchical attention networks for text classification with an emphasis on biased news. LU-CS-EX 2018–21
27. Atefeh F, Khreich W (2015) A survey of techniques for event detection in twitter. Comput Intell 1:132–164
28. Osborne M, Moran S, McCreadie R et al. (2014) Real-time detection tracking. monitoring of automatically discovered events in social media. In: Proceedings of the 52nd annual meeting of the association for computational linguistics. Association for Computational Linguistics
29. McMinn AJ, Jose JM (2015) Real-time entity-based event detection for Twitter. In: Mothe J, Savoy J, Kamps J et al (eds) Experimental IR meets multilinguality, multimodality, and interaction: 6th international conference of the CLEF association (CLEF ’15). Springer, Berlin, pp 65–77
30. Griggs JR, Grinstead CM, Guichard DR (1988) The number of maximal independent sets in a connected graph. Discret Math 68:211–220
31. Johnson DS, Yannakakis M, Papadimitriou CH (1988) On generating all maximal independent sets. Inf Process Lett 27:119–123
32. Weng J, Lee BS (2011) Event detection in twitter. InFifth international AAAI conference on weblogs and social media
33. Cordeiro M (2012) Twitter event detection: combining wavelet analysis and topic inference summarization. In Doctoral symposium on informatics engineering, pp. 11–16
34. Weiler A, Grossniklaus M, Scholl MH (2016) An evaluation of the run-time and task-based performance of event detection techniques for Twitter. Inf Syst 62:207–219
35. Cui W, Wang P, Du Y, Chen X, Guo D, Li J, Zhou Y (2017) An algorithm for event detection based on social media data. Neurocomputing 254:53–58
36. Adedoyin-Olowe M, Gaber MM, Dancausa CM, Stahl F, Gomes JB (2016) A rule dynamics approach to event detection in twitter with its application to sports and politics. Expert Syst Appl 55:351–360
37. Nguyen DT, Jung JE (2017) Real-time event detection for online behavioral analysis of big social data. Futur Gener Comput Syst 66:137–145
38. Rezaei Z, Komleh HE, Eslami B (2019) Event detection in Twitter Big Data by virtual backbone deep learning. In International Congress on high-performance computing and Big Data Analysis. Springer, Cham, pp 18–31

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

View publication stats