LITERATURE SURVEY

TOPIC: Twitter breaking news detector in the 2018 Brazilian presidential election using word embeddings and convolutional neural.

## INTRODUCTION:

Twitter has been used as one of the communication channels for spreading breaking news. We propose a method to collect, group, rank and track breaking news in Twitter. Since short length messages make similarity comparison difficult, we boost scores on proper nouns to improve the grouping results. Each group is ranked based on popularity and reliability factors. Current detection method is limited to facts part of messages. We developed an application called “Hotstream” based on the proposed method. Users can discover breaking news from the Twitter timeline. Each story is provided with the information of message originator, story development and activity chart. This provides a convenient way for people to follow breaking news and stay informed with real-time updates.

### ABSTRACT:

This is about detecting the breaking news from Twitter with the help of sentiment analysis by tracking positive, negative, and neutral thoughts. First, the detector collects the thoughts into word embeddings using natural language processing(NLP) which is a representation of similar meanings by FastText algorithm.The sentiment analysis is done by a convolutional neural network called as TextCNN which is a special tool for performing the NLP.(NLP is concerned with giving computers the ability to understand the text and spoken words in much the same way human beings can.As the words are tracked according to the emotions they are graphed and the accuracy of a particular thoughts are driven .By tracking the breaking news on Twitter for a presidential election the accuracy of 74% of the news were true and made the president to win on 2018 presidential election.

### ADVANTAGE:

* Own Your Channels.
* Better User Experience.
* Higher Engagement.
* Push notifications.
* Revenue Opportunities.
* App Store Presence.

### DISADVANTAGE:

* To maintain consistent presence of platform.

Topic:Tracking aspects in news documents.

### INTRODUCTION:

We have presented an idea in this paper for detecting and tracking topics from news articles. Topic detection and tracking are used in text mining process. From data which are unstructured in text mining we extracts previously unknown and useful information. The main purpose of this paper is to identify and follow tasks occurred in different news sources. We are going to use agglomerative clustering based on average linkage for detecting the topics, calculate the similarity of topics using cosine similarity and KNN classifier for tracking the topics.

### ABSTRACT:

In this investigation, we discuss aspect tracking, i.e., how to identify tracking story-lines of document topics. Because a huge amount of fragment information occurs, it is difficult to see what they mean and how they go within topics by hand. Here we are addressing this type of problem using stochastic models. Our basic idea is that we consider state transitions as internal structure of stories based on HMM and we extract some story-lines as aspects of topics by probabilistic likelihood. We will use KL divergence to screen topics. Also, we discuss topic detection, i.e., how to construct clusters of events to discuss the same topic. Since most of the information are textual, we should examine which words spread over which documents with huge combination. Topic tracking is to detect documents (in chronological sequences) suitable for a known topic. By using HMM, we track story-lines in documents along with temporal order, here we involve forward and backward algorithm of the top-K maximum likelihood to explore better candidates.

### ADVANTAGE:

* Get insight in to the performance of your business
* News or media coverage
* Promote awareness

### DISADVANTAGE:

* Graphics and photos do not always

appear.

* Tracking subscribers is difficult.

TOPIC: An Approach to News Event Detection and Tracking Based on Stream of Online News.

### INTRODUCTION:

A massive amount of information is now stored and exchanged via the World

Wide Web, mostly in the form of free text. Therefore, natural language pro-

cessing technologies which can map free text into structured data formats are

becoming paramount. In particular, the proliferation of electronic news media

has led to an emergence of publicly accessible news aggregation systems (e.g.,

Google News, Yahoo! News, SiloBreaker, NewsTin and DayLife) on the web for

facilitating navigation through news broadcast daily worldwide. Such news ag-

gregation systems group topically related articles into clusters and classify them

according to various predefined criteria. Although such systems provide a more

structured view of what is happening in the world, the amount of data requiring

processing by a human remains enormous. Recently, several endeavours towards developing real-time news event extraction systems have been attempted, to de-

tect key information about events from various electronic news media and sum-

marize this information in the form of database-like structures. In this way, such

systems have been successful in providing even more compact event descriptions

that combine information from different sources.

This article gives an overview of the multilingual event-extraction system de-

veloped at the Joint Research Centre of the European Commission for extracting

violent and natural disaster event information from on-line news articles collected

through the Internet with the Europe Media Monitor (EMM) [6,5], a web based

news aggregation system, which regularly checks for updates of news articles

across multiple sites in different languages. Gathering information about crisis-

related events over time is an important task for better understanding conflicts

and for developing global monitoring systems for automatic detection of pre-

cursors for threats in the fields of conflict and health. In particular, web news

reflect trends and behaviours, which constitute a powerful data source for future

event prediction. The increase in security concerns since 9/11 especially those

related to terrorism have significantly boosted research on developing systems

for automated event extraction from news.

### ABSTRACT:

Once an event occurs, usually there are a large number of online news to be released. How to quickly and accurately detect the hot events from the huge amount of online news is the focus and hotspot. Event detection and tracking technology is as a key technology to solve this problem. In this paper, we propose an approach to detect hot events from the online news stream in a timely manner and track the hot events. Based on the idea of single-pass clustering algorithm, this approach address the weight of keywords and proposes a new method to calculate similarity among news to track event. Through the analysis of the experimental results, we can find that this algorithm has a good effect on hot event detection.Based on the analysis of some news, we find that the title and text of news reports contain the keywords related to the news, can basically reflect the subject of the news reports we can use these keywords to describe this news report.

### ADVANTAGE:

* High-Quality Streams.
* Greater Audience Potential.
* Wide Variety of Content Use.
* Ease and Convenience.
* Analytics Tracking.
* Customer Support.
* Mobile Streaming.

### DISADVANTAGE:

* Require data/wifi to get online.
* Companies not making as much money due to free reading for audiences.
* News spreads quicker online - people find out news before they should.
* Lose money - can't get people to pay for digital.
* Older audiences may not access digital platforms.
* Costly to maintain.

TOPIC:Research on Topic Detection and Tracking for Online News Texts.

### INTRODUCTION:

With the rapid development and popularization of Internet technology, how to obtain useful information from massive data has become a common concerned problem. Text mining technology [1], [2] can extract effective, useful and valuable information from a large number of texts, and it has gradually become one of the key technologies to solve the problem of topic discovery. Effective analysis of massive information on the network has also become a key research content by researchers in the field of machine learning and data mining. In text mining technology, the traditional text representation method usually adopts the space vector model (VSM), which is a commonly used model in natural language processing. But this model does not take the relationship of the underlying semantics of the texts into account. The topic model is a modeling method for extracting implicit topics from massive texts. It is widely used to mine topics from documents.

### ABSTRACT:

With the rapid development of the Internet, the amount of data has grown exponentially. On the one hand, the accumulation of big data provides the basic support for artificial intelligence. On the other hand, in the face of such huge data information, how to extract the knowledge of interest from it has become a matter of general concern. Topic tracking can help people to explore the process of topic development from the huge and complex network texts information. By effectively organizing large-scale news documents, a method for the evolution of news topics over time is proposed in this paper to realize the tracking and evolution of topics in the news text set. First, the LDA (latent Dirichlet allocation) model is used to extract topics from news texts and the Gibbs Sampling method is used to speculate parameters. The topic mining using the K-means method is compared to highlight the advantages of using LDA for topic discovery. Second, the improved single-pass algorithm is used to track news topics. The JS (Jensen–Shannon) divergence is used to measure the topic similarity, and the time decay function is introduced to improve the similarity between topics with the similar time. Finally, the strength of the news topic and the content change of the topic in different time windows are analyzed. The experiments show that the proposed method can effectively detect and track the topic and clearly reflect the trend of topic evolution.

### ADVANTAGE:

* This system takes co occurence of terms into account which gives best result.
* This system will help web users to easily search information for particular topic.
* Web users will get information quickly for respective topic their searching for.

### DISADVANTAGE:

* This system extracts words rather than phrases.
* If system extracts phrases topic detection will be faster.

TOPIC: RACING THE FACE OF FAKE NEWS USING A SCLABLE BLOCKCHAIN DISTRIBUTED NETWORK.

### INTRODUCTION:

Information veracity always affects the society, whether it

comes from social media, print media or news channels. The

information that has no real facts or evidence behind it but is

presented in a way that seems accurate and is often consumed

by millions through social networking websites, television, and

other types of digital media. This type of false information is

defined as fake news. The effect of false information spread

on social networking websites is significantly high, and it has

the potential to cause disorder in society within hours for

millions of users. Such fake news propagation has the power to

change election results, spread hatred in society, affect voting

patterns, stock values and much more. The biggest tragedy

is, once the fake news becomes viral, its hard to identify the

source of origin and thus, stop its further propagation. As a

side effect, the common man today is losing his/her trust in

media and sometimes even news channels because of the lack

of proper reference checking to verify facts. Nowadays, a lot

of digital content is being published in the form of blogs,

videos, images, etc. Anyone can freely share any information

or news over social networking websites such as Facebook,

Twitter, Instagram, LinkedIn, without any actual fact checking.People can easily believe or get influenced by such fake news

and change their perception about the subject of such news

which could be about a specific community, socio-economic

practices, religion, and a given individual. Fake news is so

powerful that it can destroy the reputation of anyone, public

figure or even an ordinary person. An intuitive solution to

mitigate these problems is to use a central authority that can

monitor such digital content and regulate information flow.

However, using such central authority kills the trust model and

privacy of decentralized social networks. Due to the distributed

and decentralized nature of blockchain technology [1], it is

strongly believed by the research community that this technol-

ogy is suitable for several areas beyond finance, including e-

voting [2], [3], healthcare [4], [5], supply chain [6] and digital

right management system [7]. In this paper, a blockchain and

watermarking based social networking framework is proposed.

Our system has the ability to trace the root or origin of

fake news which will help in refraining the propagation of

fake news on social networks. In the proposed framework,

blockchain stores each news item shared or uploaded on

the proposed social networking platform in the form of a

transaction performed by registered users. Because of the

transparent and traceable nature of blockchain, it is possible to

verify the source of any information that is shared on such a

platform. Tracing the news source by using blockchain can

be achieved with the help of timestamping and the chain

connection between blocks. In order to identify the news path

shared by users on such social media platform, it is necessary

to trace news items by going backwards step-by-step, to

identify which user originated or modified the news with

malicious intent. Block headers in a blockchain contain lots of

information, e.g. a pre-block hash value, current block hash,

timestamps etc. These block headers can provide assistance

in data tracing. Once the user uploads any information on a

social networking website, that transaction can be stored on

the blockchain, and every time a new user shares that data

or attempts to modify the data, it makes a traceable chain of

transactions. With the help of timestamps, anyone can identify

the sequence of such transactions stored in the blockchain.

### ABSTRACT:

### In the news industry, as well as in social media, fake news detection and identification of news sources has become a central topic of discussion. In the era of digitization, anyone can easily generate or manipulate digital content and publish them on social media websites. On the one hand, these social networking platforms provide ample ease in modern-day communication but on the other hand, using such platforms has posed new challenges to real-world implementation like viral spreading of false/fake information with malicious intentions. In this paper, a naive block chain and watermarking based social media framework is proposed to control the fake news propagation. We postulate a new block chain model to mitigate existing challenges in this field. Moreover, the novel solution can help in reducing the spread of fake news by tracing the root or origin of the fake news on social media. Through our experimental results, we show that our blockchain-based solution is able to immediately stream data through a bloXroute server that can propagate data up to 100 times faster than conventional solutions.

### ADVANTAGE:

* Advertisers take the Advantages of Fake News
* Influencers also take benefits of Fake News
* Political Warfare
* Fun and Entertainment

### DISADVANTAGE:

* Change in Public Opinion
* False Perception
* Fake News may lead to Social Unrest
* Fake News Cost lives