

Contraversy of News with Twitter Comments

Dr. S. B. Thorat¹ and Aniket Dixit²

¹Director ITM College, Nanded, {Phd, ME, MBA, MSc, AMIE, LM-ISTE, Fellow IETC

²Research Scholer, MCA

ABSTRACT

Streaming Social media sites and apps of it , including Facebook, YouTube, Twitter and blogs, have turn out to be most important social media affection at present era. The huge amount of information from this medium has become an attractive resource for organizations to monitor the opinions of users, and therefore ,it is receiving a lot of attention in the field of sentiment analysis In this work we are analyzed some news that are having more trending characteristic due to affection day today life of common man. we are using a quantity of part of twitter comments of some duration and glance the polarity of it. With the assist of sentiment analysis . in streaming social media of different Twitter datasets(i.e. NEWS) The obtained results indicate the potential of our approach to detect trending rumor news.

Keywords: Sentiment analysis, polarity, natural language processing (NLP)

INTRODUCTION

With the rapid development of the Internet, an increasing number of people express their opinions and views social networks, and so forth) by publishing blogs and posting comments. Millions of pages of data are posted on the Internet every day, and the Internet has grown to be one of the most important sources for people to acquire information and make decisions. The importance of discovering the sentiment orientation of user-generated content has been recognized in a wide range of application domains, e.g., to help manufacturers collect and predict the attitudes of consumers toward their products and to help political associations understand the opinions of the general public. Hence, sentiment analysis is of great theoretical and practical significance, and it has attracted increasing attention from both academia and industry. Sentiment analysis aims to capture the diverse sentiment information expressed by authors in given natural language texts, and it has been a core research topic

In the past decade, the proliferation of social media services has drastically changed the news dissemination process: along with traditional news sources (television channels and newspapers), an ever-increasing audience turns to social media platforms, such as Twitter, for real-time information sharing. However, the same features (speed and ease of information sharing) that make social media a powerful dissemination tool, have also created a prolific environment for the fast spread of rumors. In psychology, rumors are defined as information or statements that cannot be verified as true or false while spreading among people. Among the many rumors that spread in social media, we focus on those emerging as popular topics and attracting tremendous social attention. We refer to them by trending rumor events. In this context, trending rumor event is defined as a group of social media posts indicating one controversial emerging topic (i.e. statement or claim), that is discussed frequently and able to reach a wide audience within a specific time period. Compared to rumors with limited attention, trending rumor events bring uncertainty into a wider audience and fuel stronger social chaos. Understanding trending rumor events is essential as they fuel mistrust, cause panic and often prompt irrational behaviors. The implications are obvious in emergency and disaster situations, when trending rumor events are categorized as fake information. For example, medical researchers are raising concerns that public health misinformation could impede efforts to limit the spread of a virus. However, the implications go beyond this. Social media platforms and Twitter in particular, provide users with the service of fast communication and information sharing. Distorted information spread through these platforms, alters the provided service as it damages its purpose of being beneficial to society, in often cases, achieving the opposite. In the light of these important implications, we focus on analyzing features of trending rumor events in social media. There are three important keywords that outline our contribution. Twitter Sentiment Analysis A popular social medium is Twitter,¹ a micro-blogging site that allows users to

write textual entries of up to 140 characters, commonly referred to as tweets. As of June 2015, Twitter has over 302 million monthly active users according to their homepage, whereof approximately 88% have their tweets freely readable. Additionally, over 84% of the users also have their location specified in their profiles [Beevolve, 2012], enabling the possibility of performing drill-down on geographic locations. Data created by Twitter is made available through Twitter's API, and represent a real time information stream of opinionated data. Tweets can be filtered both by location and the time they were published. This has paved the way for a new sub-field of SA: Twitter sentiment analysis (TSA). Performing natural language processing on textual data from Twitter presents new challenge because of the informal nature of this data. Tweets often contain misspellings, and the constrictive limit of 140 characters encourages slang and abbreviations. Unconventional linguistic means are also used, such as capitalization or elongation of words to show emphasis. Additionally, tweets contain special features like emoticons and hashtags that may have an analytical value. Hashtags are labels used for search and categorization, and are included in the text preceded by a "#". Emoticons are expressions of emotion, and can either be written as a string of characters e.g., ":-)", or as a Unicode symbol. Finally, if a tweet is a reply or is directed to another Twitter user, mentions can be used by pretending a username with "@".

Web mining is the application of data mining techniques to extract knowledge from web data, i.e. web content, web structure, and web usage data.

--The attention paid to web mining, in research, software industry, and web based organization, has led to the accumulation of important experience. It is our goal in this chapter to capture them in a systematic manner, and identify directions for future research.

--The demand for information on opinions and sentiment

"What other people think" has always been an important piece of information for most of us during the decision-making process. Long before awareness of the World Wide Web became widespread, many of us asked our friends to recommend an auto mechanic or to explain who they were planning to vote for in local elections, requested reference letters regarding job applicants from colleagues, or consulted Consumer Reports to decide what dishwasher to buy. But the Internet and the Web have now (among other things) made it possible to find out about the opinions and experiences of those in the vast pool of people that are neither our personal acquaintances nor well-known professional critics — that is, people we have never heard of. And conversely, more and more people are making their opinions available to strangers via the Internet. Indeed, according to two surveys of more than 2000 American adults each

- 81% of Internet users (or 60% of Americans) have done online research on a product at least once; • 20% (15% of all Americans) do so on a typical day; • among readers of online reviews of restaurants, hotels, and various services (e.g., travel agencies or doctors), between 73% and 87% report that reviews had a significant influence on their purchase;1 • consumers report being willing to pay from 20% to 99% more for a 5-star-rated item than a 4-star-rated item (the variance stems from what type of item or service is considered); • 32% have provided a rating on a product, service, or person via an online ratings system, and 30% (including 18% of online senior citizens) have posted an online comment or review regarding a product or service .

2. BACKGROUND

Gathering this information from social media will be one-sided; boisterous what's more, has numerous different issues; however, the impact of these issues can be diminished by taking information on a solitary topic from various sources and distinctive groups, posts, comments. We join the analysis of twitter and Facebook user communications alongside content composed by writers from various sources one-sided towards diverse topics. In this paper, we investigate the events related to government policy GST and Demonetization and news "Facebook made mistakes on Cambridge Analytica, says Zuckerberg", "kathua assault case", "A ridiculous canvas: on Padmavati", "spot settling in IPL" and "oils pills catastrophe in world". We find that our quantitative analysis can accurately arrange the topic as being controversial or on the other hand not. Numerous past studies have investigated a few parts of these posts and comments or other

social media news bolster, such as the greatest number of preferences and generally re -tweets. Generally, past works can be described as contextual investigations, where controversy is distinguished in a solitary painstakingly curate dataset, gathered utilizing plentiful space learning and helper area particular sources (e.g., a broad rundown of hash tags (#) in regard to a noteworthy political occasion). We mean to beat those impediments. Our goal is to identify controversy regarding topics without prior domain-specific knowledge about the topic in question. In addition, we aim at complementing these results by sentiment analysis on the same topic using user opinions as well as the journalist's Articles.

News or Event	Number of Tweets
GST in India	1194
Demonetization in India	38
Facebook made mistakes on Cambridge Analytica, says Zuckerberg	109
Kathua assault case	974
Oils pills catastrophe in world	75
A ridiculous canvas: on Padmavati	3977
Spot settling and Fixing in IPL	15443

Table: Dataset of Comments

The news media or news industry are forms of mass media that focus on delivering news to the general public or a target public. These include print media (newspapers, newsmagazines), broadcast news (radio and television), and more recently the Internet (online newspapers, news blogs, etc.) Every newspaper publish online nowadays. They has their own websites for publication, the editors of newspaper publishes an editorial article on many news or events which provide accurate and important information about news and events .it helps to find out controversy of news and event .We fetch editorial articles, blogs by scientist and researchers regarding news or event from different standard newspapers, online scientific blogs, websites and magazine.

Our approach

analysis on article form a train and testing data sets for machine learning approach ,we test on both approaches of the sentiment analysis, at last we propose new mathematical model to evaluate controversy score .We find the CS of Comments and CS of Sentences using formula number 5 if the percentage of negative comments or sentence is less than or equal to 50% ,6 if the percentage of negative comments or sentences is greater than 50%.After the getting values of CS of comments and CS of Sentences we take Average of both to calculate the Final Controversial Score

$$\text{I.e. FCS} = (\text{CSC} + \text{CSS})/2$$

Where,

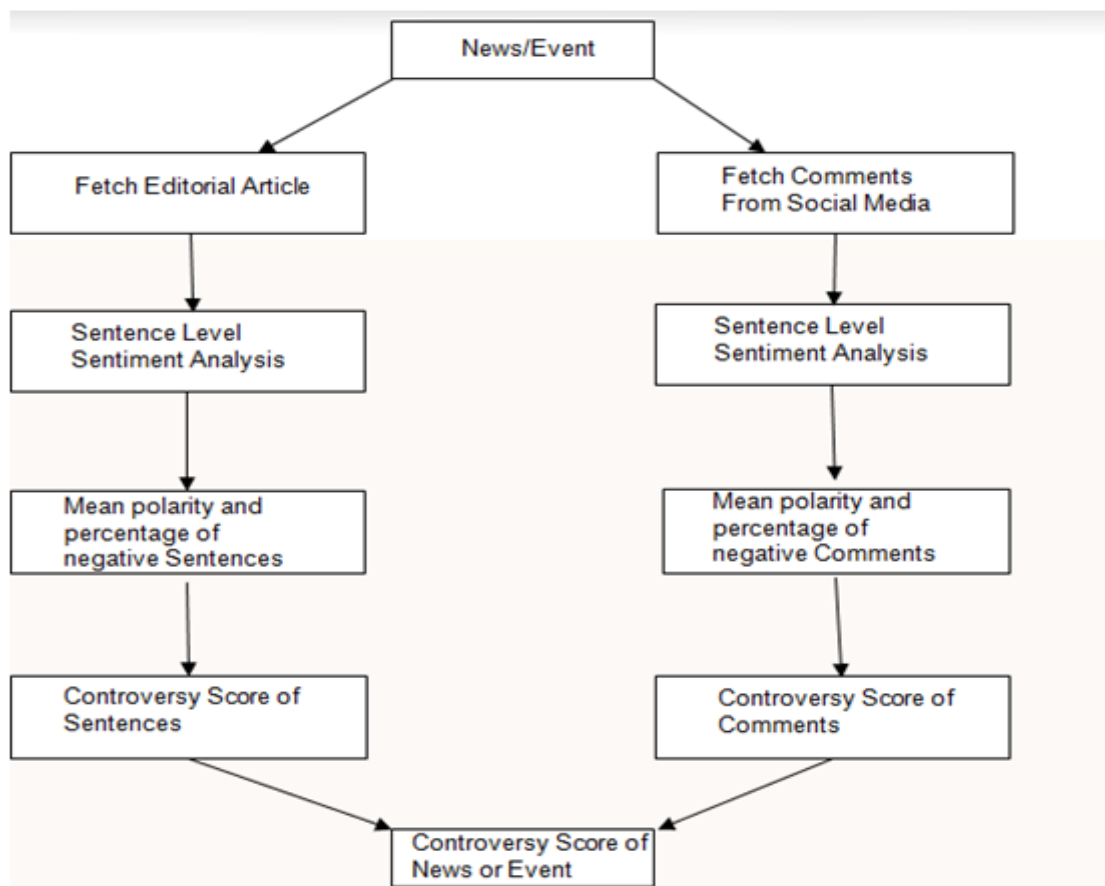
FCS = final controversy Score of the News or Event

CSC = Controversy Score of Comments

CSS = Controversy Score of Sentences

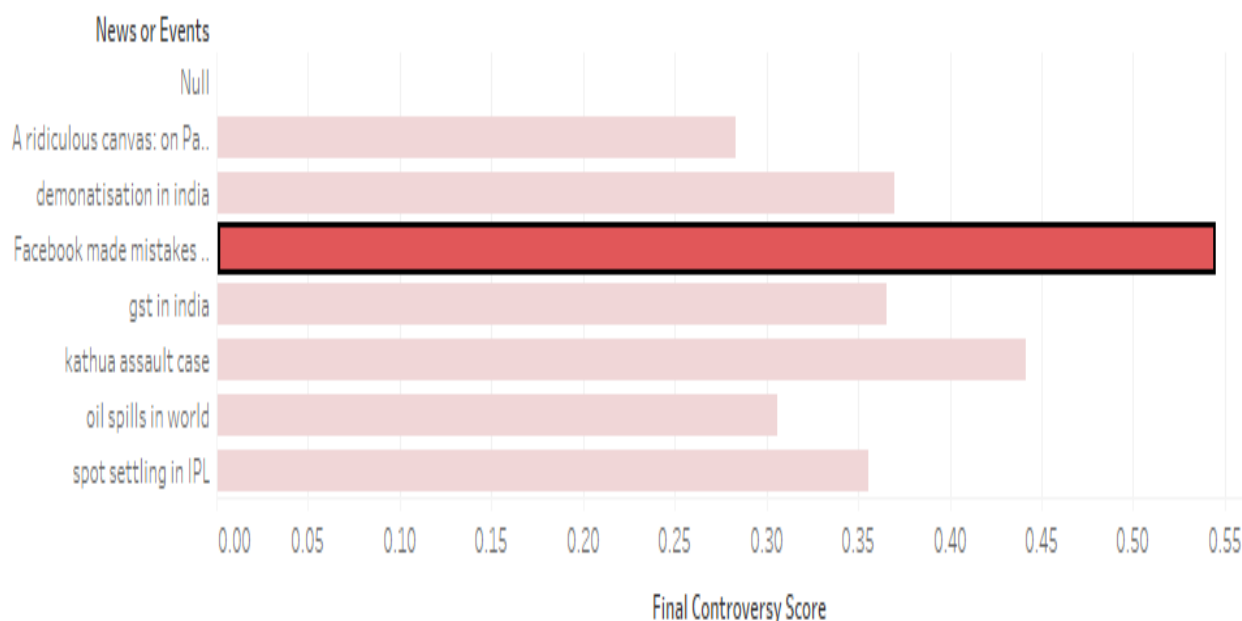
Here question arises that what will be the value of proportionality constant k

To determine the value of k we propose a table method or polarity percentage method.



News_or_E vents	Percentage_ of_Negative _Comments	Mean_of_N egative_Pol arity_Com ments	Percentage_ of_Negative _Sentences	Mean_Nega tive_Polarit y_Sentences	Controve rsy_Score _Comme nts	Controv ersy_Sco re_Sente nces	Final_Contro versy_Score
GST in india	24.390244	-0.30204	18.70504	-0.1183	0.165254	0.278527	0.365365579
demonatisati on in india	28.947368	-0.28409	18.87755	-0.16698	0.468716	0.271332	0.370023744
Facebook made mistakes on Cambridge Analytica, says Zuckerberg	44.036697	-0.11633	17.94872	-0.13055	0.757074	0.329974	0.543524252
kathua assault case	23.613963	-0.21823	27.38095	-0.27333	0.422	0.46081	0.441404777
oil spills in world	21.333333	-0.18681	22.47191	-0.16	0.274082	0.337082	0.30558174
A ridiculous canvas: on Padmavati	13.050038	-0.24766	21.4876	-0.23226	0.205501	0.360814	0.283157361
spot settling in IPL	16.849058	-0.26522	20.89552	-0.1081	0.24776	0.463907	0.355833502

Table: ALL CONROVERSY NEWS RESULT



CONCLUSION AND FUTURE WORK

In this study we analyzed social network activity on the different topics and on popular news sources from different categories with a special focus on polarity of the user interaction. We performed the first combined study of sentiment and textual analysis for quantifying controversy in social media. Our results suggests that user inter action from all these different categories of news and events will give the overall views or opinions of the people on that particular topic which really contributes to the classification of a topic being controversial or not. Besides, our process is domain Independent and can be applied to any dataset from other domains. From the application point of view, our controversy score can be used to generate recommendations for trending or hot topics in news feeds on social media. With the above bar and line graphical output of table shows in comments of social media facebook get higher range and in sentence level kathua assault case and IPL spot fixing get level equal but in ranking of final news controversy i.e. final controversy score is higher ranked by Facebook case

REFERENCES

- [1] Sentiment analysis and opinion mining a book by bingliu
- [2] IEEE paper by author Allaparthi Sriteja Controversy Detection Using Reactions on SocialMedia
- [3] IEEE paper by author Ana-Maria Popescu, Marco Pennacchiotti Detecting Controversial Events from Twitter
- [4] IEEE paper by author Efthymios Kouloumpis, TheresaWilson, Johanna Moore Twitter Sentiment Analysis: The Good the Bad and the OMG!
- [5] Springer paper by authors Lu Wang, Claire CardieA Piece of My Mind: A Sentiment Analysis Approachfor Online Dispute
- [6]Book by Author Steve Y. Yang¹, Sheung Yin Kevin Mo² Social media and news sentiment analysis for advanced investment strategies
- [7] A Research paper by Kazimierz Zielinski, RadoslawNielek, Adam Wierzbicki, Adam JatowtComputing controversy: Formal model and algorithms for detecting controversy on Wikipedia and in search queries