

## Controversy News Ranking With Social Media

**Dr.S.B.Thorat**

Director Itm College, Nanded

M.C.A.

(Phd,Me,Mba,Msc,Amie,Lm-Iste,Fellow IETC)

**Aniket Dixit**

(Research Student)

### Abstract—

The use of online social media is expanding; numerous political gatherings, associations and news sources are utilizing this as a stage for spreading news. This causes us to gather more information about individuals opinions on a topic and use this information as the way to see a few perspectives regarding controversy. In the current circumstances, social media has been assuming a part in how people process data and shape opinions. In this way, there has been a need to offer the user a view that varies from what they are for the most part presented to, for picking up the in general picture on a topic before arriving at a conclusion or an assessment. Be that as it may, behind every single such application, the crucial yet testing assignment that should be fathomed is to naturally order whether a topic of exchange is controversial or not. As well as rank by the news with final controversy score

**Index terms--**Controversy Detection, Social Media Mining, Sentiment Analysis,

### Introduction

On these days and age the irrefutable fact is that social media users are extra likely to share how they sense concerning a present “hot topic” on social media platforms. Hot topics are present relationships that may trend regionally or internationally. Users may post negative, positive or neutral .opinions about that topic or a exacting product they are using. The progress of artificial intelligence (AI) has opened doors in which we can write down algorithms to facilitate users notice and classify online opinions. This study aims at proposing an AI model to detect emotion in unstructured texts. We examine the sentiments of user views about the recently controversial issues and calculate up to them with the associated trendy topics. The main participation is to power on social media to approximation a sentimental opinion appraisal on the most modern trends or topics of controversy on Twitter. Our main objective is to get hold of perceptions of the users’ opinions based on the number of likes, retweets and using a natural language processing (NLP) toolkit to determine the sentiment of texts. Experimental results verify that sentiment analysis is precious

to identify users’ likes, comments, and retweets on a product

The current outbreak and widespread of microblogging and social networking websites are reshaping many aspects of the modern day social interaction. Social Media today has exceeded the limits of entertainment or simple social interaction contexts. Social media now is better described as a living organism, that has a structure and a soul. The Social Media now reflects the pulse of the people; it reacts to their emotions, and interacts with their opinions. Analyzing and monitoring the content of Social Media can bring about some valuable insights, of which the conventional media means weren’t able to convey. This project tackles the concept of community detection within Social Media means, Twitter in particular. The virtual communities that the people on Social Media tend to group themselves into can provide precious sources of information regarding the patterns of communication and knowledge propagation. The project identifies the dimensions of similarity and interaction between any pair of users and provides the tools to calculate them. These dimensions are used as building blocks to construct the network structure to be used to detect the communities.

The project presents a number of algorithms and software systems implementations that approach the issue of community detection in Twitter. The parameters and factors affecting the community detection process are then investigated in depth, to help tuning the procedure to provide the most efficient results. The problem of sentiment analysis is then studied in the context of the detected communities. The project implements a procedure in which the detected communities are utilised to enhance the overall performance of the sentiment analysis process.

**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.

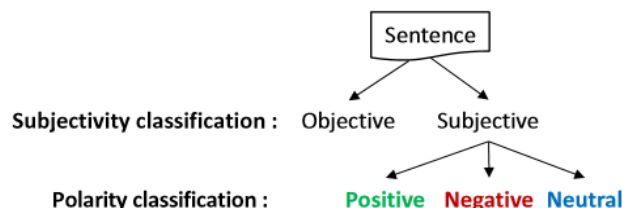


Figure- sentence analysis

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, news magazines), broadcast news (radio and television), and more recently the Internet (online newspapers, news blogs, etc.) Every newspaper publishes online nowadays. They have 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:**

To overcome the limitation of the old Controversy Detection Approaches we proposed a new Controversy Detection Approach, in this we define a new framework, in second part we explain some data pre-processing techniques then we write about feature extraction for sentiment analysis, calculating the polarity of the comments, use word level sentiment analysis on comments and sentence level sentiment 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=finalcontroversyScoreoftheNewsorEvent

CSC=

Contro

versyS

coreof

Comm

ents

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.

Following algo shows our approach

Algorithm Controversy Detection

//Input and output

Input : A (Articles), T (Topic),t(tweets),

Consumer\_Key, Consumer\_Secret,

Access\_Token,Aaccess\_Token\_Secret

Output : Controversy Score

//steps

while article in A, tweet in t do

  r ← reactions

  s ← share

  c ← comments

  r a ← average reactions

  s a ← average share

  c a ← average comments

  sent ← sentences

con\_sent ← article.sentences with controversial vocabulary

com\_sent ← Total sentences in all comments of an article

com\_con\_sent ← sentences in comments with controversial vocabulary of an article

tweet\_sent ← Total sentences in all tweets of topic T

tweet\_con\_sent ← sentences in tweesys with controversial vocabulary of topic T

pos ← positive comments

neg ← negative comments

i ← (r/r a + s/s a + c/c a )/3

l←(con\_sent/sent)+ (com\_con\_sent/com\_sent)

+(tweet\_con\_sent/tweet\_sent)

if 1.5 pos<neg then

  sentiment\_article ← (pos - neg)/(pos + neg)

sentiment\_tweet ← (pos - neg)/(pos + neg)

else sentiment\_article ← 1 - (pos - neg)/(pos + neg)

sentiment\_tweet ← 1 - (pos - neg)/(pos + neg)

end if

controversy\_score\_article ← i \* l \* sentiment\_article

controversy\_score\_tweet ← i \* l \* sentiment\_tweet

if controversy\_score\_article > 0.7 then

  article.label ← controversial

  final\_score\_article ← final\_score\_article + 1

  article.label ← non controversial

end if

if controversy\_score\_tweet > 0.7 then tweet ←

  controversial

  final\_score\_tweet ← final\_score\_tweet + 1

  else

  tweet ← non controversial

  end if end while

iffinal\_score\_article/len(A) > 0.5 then T.label ←

  controversial

  else

  T.label ← non controversial

  end if

iffinal\_score\_tweet/len(t) > 0.5 then

  t.tweet ← controversial

  else

  t.tweet ← non controversial

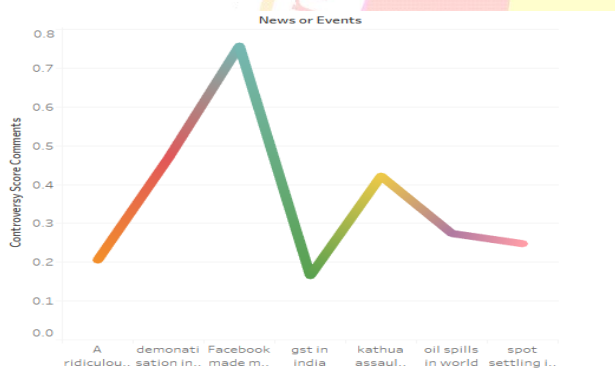
  end if

  return T.label, t.tweet

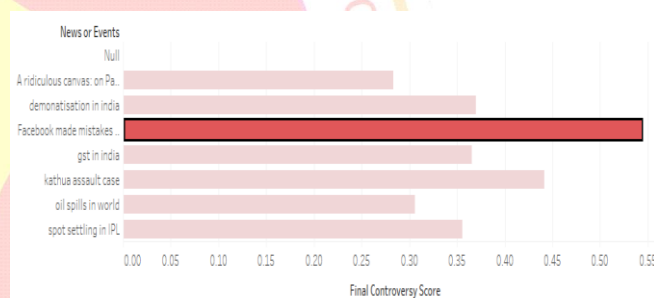
**Table:ALL CONROVERSY NEWS REESULT**

News_or_Events	Percentage_of_Negative_Comments	Mean_of_Negative_Polarity_Co_mments	Percentage_of_Negative_Sentences	Mean_Negative_Polarity_Sentences	Controversy_Score_Comments	Controversy_Score_Sentences	Final_Controversy_Score
GST in india	24.390244	-0.30204	18.70504	-0.1183	0.165254	0.278527	0.365365579
Demonatisation 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

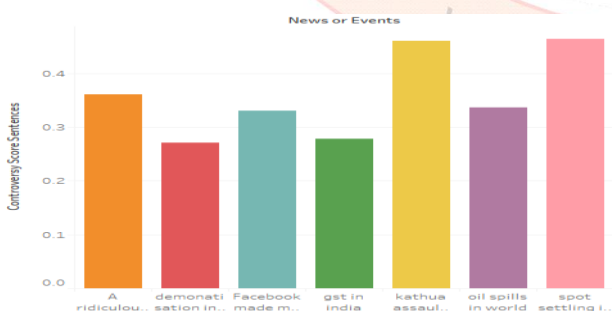
**GRAPH:CONTROVERSY SCORE COMMENTS**



**GRAPH3:FINAL CONTROVERSY SCORE**



**GRAPH2:CONTROVERSY SCORE SENTENCES**



**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

In future work we intend to investigate these metrics in other domains to generate a generalized platform to give details with respect to controversy and extend this work to find predict the probability of a post being fake or not.

### References

1. Sentiment analysis and opinion mining a book by bingliu
2. IEEE paper by author Allaparthi Sriteja Controversy Detection Using Reactions on Social Media
3. IEEE paper by author Ana-Maria Popescu, Marco Pennacchiotti Detecting Controversial Events from Twitter
4. IEEE paper by author Efthymios Kouloumpis, Theresa Wilson,
5. Johanna Moore Twitter Sentiment Analysis: The Good the Bad and the OMG!
6. Springer paper by authors Lu Wang, Claire Cardie A Piece of My Mind: A Sentiment Analysis Approach for Online Dispute
7. Book by Author Steve Y. Yang<sup>1</sup>, Sheung Yin Kevin Mo<sup>2</sup> Social media and news sentiment analysis for advanced investment strategies
8. A Research paper by Kazimierz Zielinski, Radoslaw Nielek, Adam Wierzbicki, Adam Jatowt Computing controversy: Formal model and algorithms for detecting controversy on Wikipedia and in search queries

