IJIST
P ISSN : 2618-1630
E ISSN : 2909-6130

Meme Detection of Journalists from Social Media by Using Data Mining Techniques

Original Article

Sajawal Khan [1] , Adeel Ashraf [2] ,Muhammad Shoaib [3], Muhammad Iftikhar [4], Imran Siddiq [2] , Muhammad Dawood Khan[5], Abdullah Faisal [2]

[1]Lecturer, Department of Computer Science, Afro Asian Institute, Lahore, Pakistan
[2]Lecturer, Department of Information Technology, Afro Asian Institute, Lahore, Pakistan
[3]Lecturer, Aspire College, Manawan Campus, Lahore, Pakistan
[4]PHP Laravel developer
[5]Lecturer Computer Science at Layyah Campus GCUF, Pakistan
* **Correspondence**: Imran Siddiq, imrancpe4u@gmail.com

Citation | Abbas. N, Nimra, Habib. W "Monitoring of Mangrove Cover of Western Indus Delta Karachi Pakistan". International Journal of Innovations in Science & Technology, Vol 04 Issue 04: pp 1055-1069, 2022.
DOI | https://doi.org/10.33411/ijist/20224410551069
Received |March 27, 2021; **Revised** | April 15, 2021; **Accepted** | April 20, 2021; **Published** | April 25, 2021.

With regard to today's social media networks, memes have become central. Share a million memes per second on different social media. The detection of memes is a very concentrated and demanding subject in the current field of research. Share about a million memes on various social media. Today's social media (What's App, Twitter, and Facebook) is widespread around the world. Since people in all countries are heavily used, use a set of social media application data (Twitter). As social media has an enormous amount of data overall the world. That will use that data set of Twitter for meme detection of journalists in Pakistan. So, this is done by getting data set from social media (Twitter, What's App, and Facebook) by using their authenticated APIs. For this analysis use some opinion mining techniques and sentiment analysis like the statistical descriptive and content analysis. In our society, it is the better way to analysis about any journalists because social media can provide very huge amounts of data about any journalist which is true or false but I can make approximate correct perceptions by using sentiment analysis and text mining techniques. It will provide highly wanted and hidden characteristics and perceptions for searchers and demanding people about journalists. Finally use for sentiment analysis by using Python.

Keywords: Meme Detection, Journalists, Social Media, Data Mining, Python, Sentiment analysiss.

Acknowledgment.
This declaration clarifies that all Authors have seen and approved the paper that has been submitted. We assure you that the article is the original work of the Authors. We ensure that the article has not been previously published and is not being considered for publication anywhere.

Author's Contribution
All authors contributed significantly to the study, and all authors agree with the content of the manuscript in IJIST.

Conflict of interest
The authors declare no conflict of interest in publishing this manuscript in IJIST.

Project details. NIL

Introduction

Memes are the thoughts or ideas of others about others. Memes are concepts or behaviors that are passed from one person to another. Examples of mimes include beliefs, fashion, stories, and sentences. In earlier generations, memes were often divided into local cultures or social groups[1]. However, just as the Internet has created a global community, it can be spread across countries and cultures around the world. Communication Online communication memes are called "internet memes". Some people try to find and expose others' mistakes rather than correct them[2].

As quoted in more than 100 publications now, biologist Richard Dawkins first described the concept of MEM in 1976 as "the unity of cultural communication or copying" (192, the original text is the focus of interest)[3]. Comparing genetic and cultural concepts Dawkins, its concept has been a source of excitement, insight, controversy, and skepticism across the various academic disciplines[4]. The term mimetic was coined by extending "mime" to its genetic experience, which became an independent field of research in the early 1980s, 1990s, and 2000s, especially with Robert Anger, Susan Blackmore, and Richard. Brody, Daniel Dennett, Kate Dustin, Aaron Lynch, and Tim Tyler have published their images on the phenomenon over the years[5]. And opinion. Although Dawkins is known as the "Guardian" of this brand, he is rarely associated with the game[6].

However, the label "Daojin Scene Memes" is appropriate for these experts to treat the subject. The scientific community has largely ignored Dawkins' memes. Dan Sperber described the similar scheme as "misleading" (172) and Peter Kinnan's "new wise words that evoke the ontology of pseudoscience"[7]. Even rhetoric expert Caroline R. Miller described Mime as inferior, and in a speech on gender development in 2012, she began discussing memes: "I really hope I don't even have a talk about memes Not to do "Although the cause of the fear is not fully understood, it is likely to be related to the widespread belief that meme is a pseudoscience, he also mentioned[8]. How the model camp of the Mimics is perceived as an attempt to explain culture through science rather than humanism, a move that has been strongly criticized in both humanity and science, for example, Good Heart. Miller also points out that as Glock argued, the same journalist lasted only eight years, which is essentially an indication that the disappearance of the leading newspapers in the area is additional proof. His cheating character [9]. Wrote an article that provides a complete literary review of mimes and ponders and questions these mimes.

Taylor thinks the word "model science" might be more appropriate. That is, rather than marking mimetic as pseudoscience because the subject (human brain) has not been studied at the level necessary to establish empirical evidence of mime[10]. Meme (assuming meme is the reality and studying its external manifestations) and intracranial memes (trying to understand how memes are made in our brain) the two opposing authors of this racism are Suzanne Blackmore and Robert Auger. Revelations detected as mimes[11]. On the other hand, Unger in Electrical Memes has a lot of hard-working scientific arguments on how memes can exist in our brains, but Nissos Rasciens still can't do that to test its theory. The difference between the two camps can be attributed to the failure of the received Hindi memes [12].

Literature Review

In the context of the federal election of [13]. About 100,000 journalists were studied. Point out that Twitter is widely used to spread journalist relevant journalist information. The number of journalists only shows accurate results, indicating that the site's Twitter information better illustrates this situation with lefts. Line online strategies can be used. Plan the outcome of the selection to some extent. Similarly[14]. Studied the use of Twitter in the Swedish elections of 2010 and found that Twitter served as a channel for promoting journalist content rather than as a channel for journalist dialogue [15].

The Internet provides information on research, gender stealing, soil scholarship, and multimodal transport studies mammography. McGee's suggestion to explain cultural actions through rhetorical analysis can be resolved through a critical analysis of dialogue. This is why this paper adopts a conceptual discourse that will be broadly based on the theory of gender and copy. This speech analytics project is dedicated to the five main mums in the chat online chat room and some related examples[16][17]. Some of the rhetorical concepts mentioned above will be explicitly analyzed in relation to the art of pantomime, while others are often mentioned. In a nutshell, I am going to study Simulacra, conduct oral assessments, and hope to provide work-related analysis and theory in progress to understand Internet simulators as complex multimodal compounds[18].HEMOS (Humor-EMOji-Slang-based) system for fine-grained sentiment classification for the Chinese language[19]. Twitter is the most authentic social media romfrom the research point of view. We get tweet different journalists from TTwitter ttheirheck theand ir popularity and apply Sentiment analysis. Sentiment analysis is the main part of my research; other researchers also apply sentiment analysis on social media (Facebook, Instagram, and YouTube). The main point is that they apply on the database not in real time [20].

Therefore, the article is a symbol of the area's biggest problem, that is, researchers are trying to define this phenomenon, but due to the wider understanding of the vocabulary, frustration with the word definition is not always encouraging[21]. Forced to create opposition. General Chat Room Chat Room 132 This article examines the use of pantomime by reading close-up of memes found in different situations, as well as analyzing the rhetorical style of meme art on the Internet, including popular journalists and sociologists. Get their conversations and discussions online[22]. This document not only clarifies the current definition of Mimes, but also the impact of the role and role of these works of art in different social environments, their tolerance to tolerance, and the implications of multi-modal transport[23-24]. This analysis will shed light on how the concepts of gender and moderation work together and learn more about the important and useful things of the 21st century.

Results and Discussion

As we all know, the use of the Internet is increasing rapidly every day. People use the Internet to find anything and can check every detail. Nowadays, before starting any job, people must first browse, check their reputation, and comment on any work, and then decide to do so. As a result, spam comments have increased very quickly. These companies hire spammers to write spam comments about products designed to improve their business, but sometimes they hire spammers to disrupt the business of other competitors. There are two types of comments, one positive and the other negative. Different researchers have studied this and tried to find different effective methods to detect these false comments.

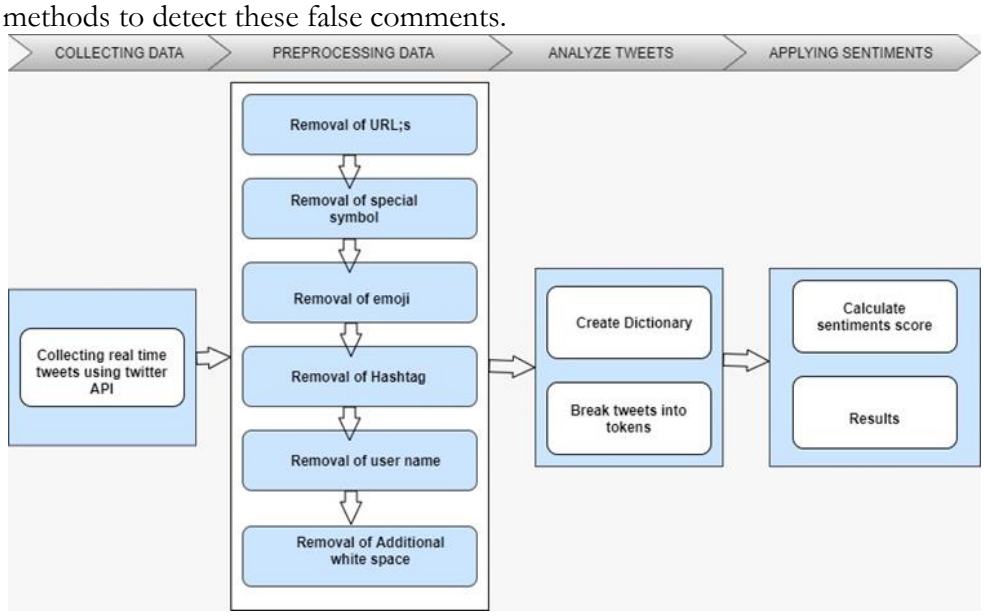

Figure: 1 Proposed Model

Some work very well and some are imperfect or require more work. Therefore, I use API technology to get the exact tweets from tweeters using a PIN. Twitter API Access Secret Key. We get the API of Tweeter by following the steps:

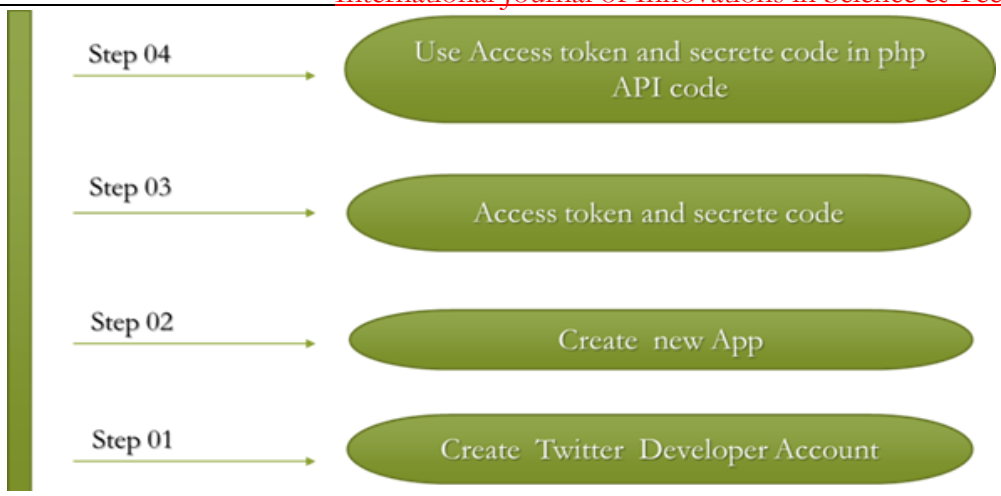

Figure 2: Access API steps.

Text Data Retrieval & Processing

For the purposes of this article, the textual content is extracted from five sources in two areas. Online reviews collected on Amazon, hotels, restaurants, and doctors, as well as tweets collected on Twitter, constitute our emotional dataset. All data sets are changed to a single, comma-separated instance. Value file (CSV) per line. These CSV files contain extracted text and emotion tags, sorted in negative or positive form. For the election dataset, the user name and location are also included in the CSV file. This file is encoded in UTF-8 format to contain most character representations. Most character representations.

Tweet Sentiment Data

The feeling corpus100 for the first tweet data set. The Emotional 100 corpus is used to create training and test sets for the analysis of feelings in tweets. The corpus includes Twitter tweets based on the presence of emojis in the text. The combination of symbols and characters is called an emoji and is used to express emotions such as express and the like. The tweeter data can be extracted via the REST API. The tweet is marked as positive or negative due to the presence of eight emoticons. Emoticons are removed from the text once the instance is marked as positive or negative. The removal of emoticons is necessary because they provide a direct relationship with the class tags and can degrade the machine-learning methods. The final dataset contains 1.6 million tweets, 800,000 positive cases, and 800,000 negative cases. The dataset has no class imbalance because it consists of a perfect 50:50 position.

Tweet text

The second Twitter-based dataset is a collection of tweets related to the word top journalist. This dataset contains tweets collected from January 22 to Nov 8, 2020, using the Twitter REST API. There is no place of execution or restriction of account, but the terms of the request are limited to subjects related to Republican and Democratic. Here are the terms of our search query: "#Shahid Masood", "#IMRAMKHAN", "#Mubasher Luqman", "#Aftab Iqbal", "#Walter Cronkite", "#Veronica Guerin". Tweets, usernames, queries, and locations are extracted into the CSV file. The dataset contains 2,942,808 tweets from 705,381 unique user accounts. No preprocessing is performed on the dataset, which means that the transferred and duplicate tweets are included in the dataset. It should be noted that this dataset does not express a tag when it is collected and adds emotion during post-processing.

Text Mining & NLP

To generate acceptable characteristics, spatial text mining is based on classical methods of NLP and machine learning. Used to change the original text on the appropriate machine. The learning function and Word-TF are very well-known technologies. Once you have changed the appropriate function space, start the machine. Training classifiers can be trained. In this section, a brief overview of the tradition is given. The feature extraction method and the machine learning classifier used in this study are given:

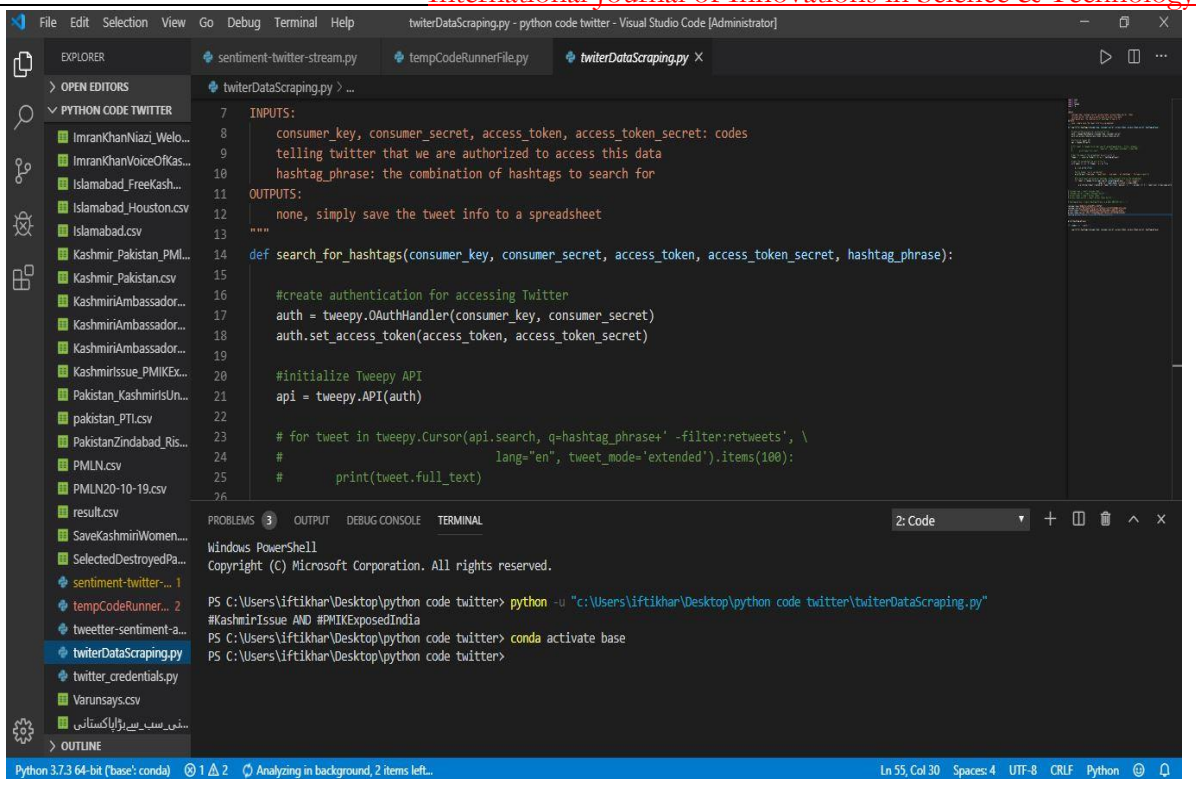

Figure 3: Creation of CSV file.

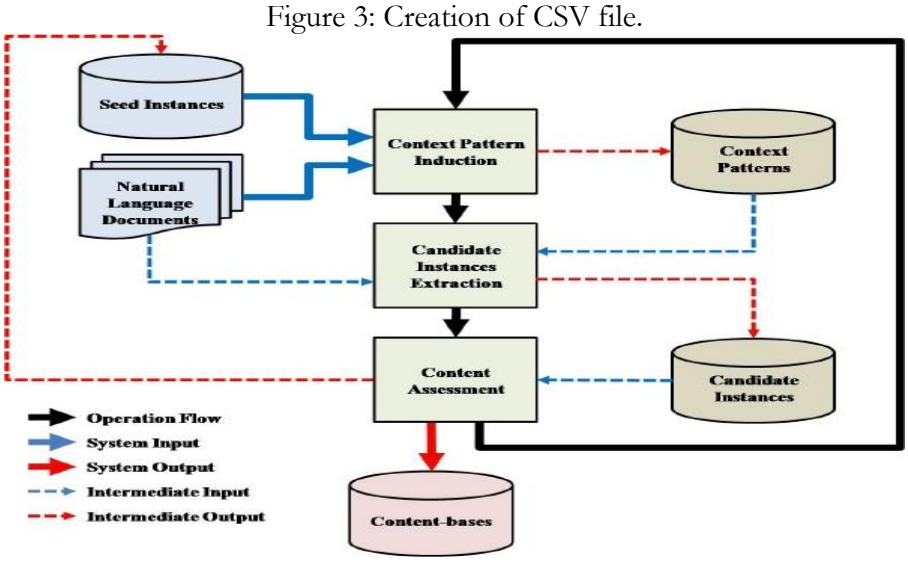

Figure 4: NLP Model

Feature Engineering & Extraction

Use the machine to perform the initial task of any form of text mining. The learning classifier is a feature. Engineering and extraction. Features engineering is an American program. Several methods are available for the function. Engineering method. The features can be extracted directly from their text using innovative NLP technology, such as part of a speech. Label, vocabulary. Cartography or grammar Characteristics The POS tag refers to the words in the document, such as verbs, nouns, adjectives, etc., depending on the corresponding part of the speech in the document. The term function is centered on the words, the words, the characters, the words "the", "of", "a", etc. Although these complex functions add to the functional space of the text, the combination of these more complex functions brings unpredictable benefits. Several machines. As a process for engineering more complex functions, the original text must be converted into a vector function for automatic learning classification. There are two popular methods to achieve this: word. Bag (BoW) and TF-IDF. (Adamic et al. 2016)The more common of the two technologies is the BoW program. The simplest form is to display the words as features. BoW uses features based on the n-gram. An engineering method in which n determines the number of words used to generate a new feature.

A boW can be implemented using a single group of words (n = 1), a group of two words (n = 2), or a group of three words (n = 3). When you use a combination of words, every single word is replaced by a function consisting of "0" or "1" which refers to the existence of the word in the document. The same process is used for two and threefold tuples, but the double words do not display the words of each other, but rather use consecutive word words. Although simple, the technical BoW is considered very practical. In addition, the feature space generated by BoW can be combined with more innovative feature engineering methods.

Cleaning

Data preprocessing is a normal first step before using neural networks to form and evaluate data. Machine. The learning algorithm is as good as the data you provide. The reliability of correctly

formatting data and maintaining meaningful functionality to achieve optimal results is serious. For computer vision computer learning algorithms, data preprocessing includes several steps, normalizes image input, and reduces size. These goals are to eliminate some of the unique features that are not important between different images. Functions such as darkness or brightness do not help to mark images. Similarly, in the task of marking the text as yes or no, the existence of a partial text does not help in any way. The pretreatment of the data is usually iterative. The tasks are not linear. Mission. This is the case in this project, where they use a new non-standardized dataset. When we discovered that neural networks were learning some worthless features, they learned some crucial pretreatment from the data.

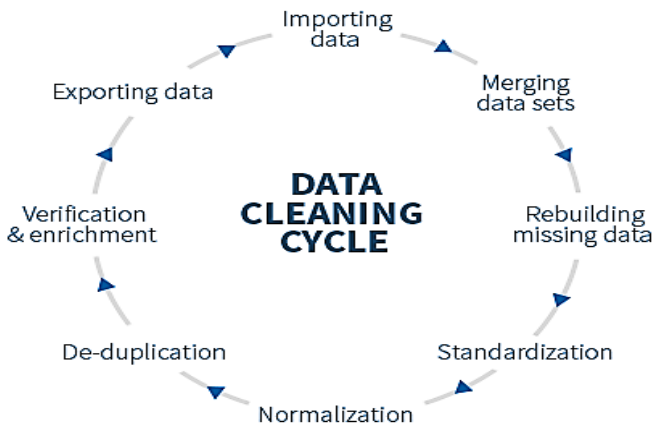

Figure 5: Data Cleaning Model

Deleting non-English words

The two results that led us to perform more pretreatment are that the following words and proper nouns exist in the most important taxonomic triads. An example of what a researcher often sees about word sequences is that the "most false" triad category is "Not My journalist ", which comes from the "tag" trend on Twitter. There are also three decisive conjunctions, such as the simple pronouns of "Aftab Iqbal". Precise names cannot help the machine. Learning. Try to detect algorithms that indicate the language pattern of news or spam significantly. They want their algorithms to be independent of the subject and make decisions based on the type of words used to explain the subject(Kleinberg et al., 2009).

Another algorithm can be designed to check the evidence in the news. In this case, it is crucial to keep the proper names/topics, as the proper name in the phrase "Donald J. Trump is the current president" has been replaced by "Hillary Clinton is the current president". Replace the classification of the real facts with false facts, but their purpose is not to examine in detail, but to check the language model, so removing the proper name helps guide the learning algorithm. Automatic in the right direction until the expression function is found.

We removed the word "non-English" using the PyEnchants form of the English dictionary. This also describes the deletion of numbers, which is unnecessary for this sorting task and this website. Although linking to a website helps to rank the page ranking of an article, it is not helpful for the tools they are trying to generate(Stieglitz & Dang-Xuan 2013).

Non-performing loan

Use the natural processing language as follows:

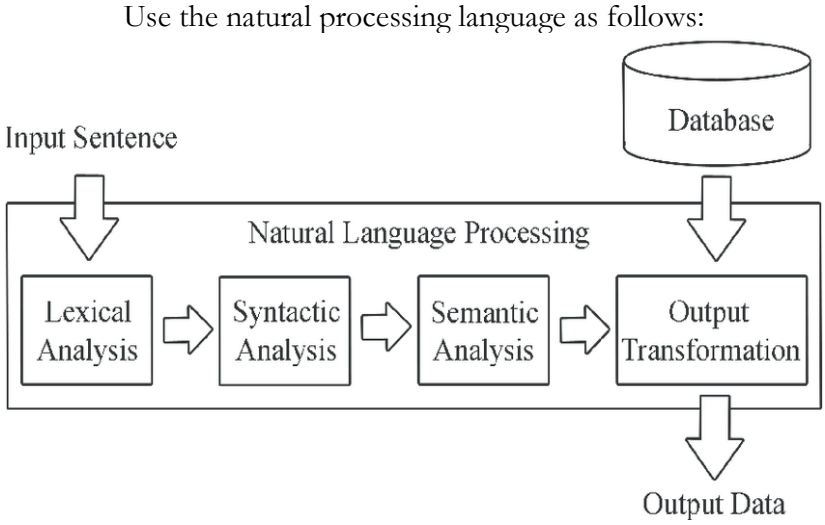

Figure 6: NLP Lexical Analysis

Sampling

During the data collection process, I created Nurses (photos, text, and text) almost February 2014, Facebook, February to April. This is equal to the time of the crisis in Crimea. In February 2014, Russia and Ukraine were concerned about the ownership of the Ukrainian islands, which caused Russian-backed forces to deploy in Crimea. Next, all questions about Crimea's independence (March 16, March 16) (Declaration of Crimea and Consequences of Land Integration

in Russian Territory).From March to April 2014, the following topics have increased several times between traditional and electronic media. After the centrifuge arrived and the data collection was completed, in May 2015, the post-Crimea debate continued for more than a year. The recent debate on the economic and journalist conflict between Crimea and subsequent Russia and the West (May 2014 to May 2015 to May 2015, 2015) has received widespread disagreement on Russian domestic and international policy. However, I decided to focus on the Crimean crisis from February to April 2014, at which time most of the journalist notes in the collection seemed to be focused on the Crimean case. I will be studying the same field at this time. The data I collected was initially filtered and 624 mummies were discussed specifically for the Crimea integration issue and distributed between February and April 2014. These main database texts constitute the first quantitative phase of the study (content analysis). The representation of this sample may be limited because it reflects my own set of issues and analyzes, but mixing it with the results of other methods can continue. In addition, my supervisor's classification review also confirmed the results because his review was independent of any experience and served the purpose of my research. Although it is difficult to prove that cumulative representatives represent all journalist debates on Russian Twitter, they still present important themes, processes, and patterns of journalist exchange during the Crimean crisis. I focus on collecting and analyzing textual images and mothers. Specifying the type of study memory is important to limit reading and to ensure that text is compared. Memes, 96, which has textured images, uses picture frames and text, making them much richer and richer than normal visual or text effects. However, when the sample is sufficiently expressive and the view text tables are involved in other topics, I also include a small number of plain text privileges in the sample. Ideal memory analysis of a large content analysis sample of 624 coded texts creates a recent text analysis sample. To be a "model", nurses must meet one or more of the following criteria: intelligible references to repetitive rhetorical concepts, a general understanding of these topics, and visual and verbal expressions. See the results (and popular quotes). A contextual analysis was performed on a sample of 50 texts. In addition, content analytics makes it possible to identify prominent Russian Twitter participants who share. I've done social network analysis on 65 users posting poems on Twitter in Crimea's best time. These accounts were the most active partners to my knowledge (each distributing at least one or more memoirs in Crimea during the monitoring period). This was a non-representative sample, but an indicator for my particular study. The meme is a strange link; it is difficult to find, so it was not possible to get a meme from the meme partners. Russian Twitter users differ in how many posts they post: some post regular texts, others become more active at certain times, while others may accidentally breast to see. This means that the user cannot be identified as a meme member for each episode, but at least one can study the role of users in a specific case study by discussing Crimea. I have segmented accounts that at least one mother explicitly shares with the pro-government as "pro-government," and another that classifies at least one of the opaque opiates as "anti-government." In my research, I used nongovernmental social network analysis to build an understanding of communication patterns among pro-government users, then internal relationships between key users, and finally explored the links between two ideological "camps". The fourth stage of my research was conducting in-depth interviews with 15 renowned Meme makers. In order to identify suitable interviewees, I monitored Twitter's journalist discussions before starting the data collection and data analysis. In addition, before starting my Ph.D., I worked for a decade as a journalist in broadcast and magazines in Russia, which introduced me to the journalist and media environment. My expertise allowed me to identify the first group of influential personalities in digital journalist dialogue. I judge their popularity by the number of followers, offline reputation (if known), and the number of retweets and followers of other popular and established accounts. Then the extra searches in their fan networks and resorts empowered me to identify more effective journalist microblogging. In addition, consulting with Russian journalists and social media experts have forced me to hire a wider and wider network of pro-government and anti-government partners in Miami. To further confirm my example, I was writing a Twitter memoir daily in July-August 2014 that warned me about topics, leading activists, and changing the tone of the conversation. The constant fragmentation of my personal meetings with media collaboration experts has led me to target the nearly 200 Crimea-based mayhem participants on Twitter. I contacted about 50 people and asked them to participate in my study; more than half did not respond and others rejected the offer. After several rounds of discussions, I was able to conduct 15 interviews with influential colleagues.

Because false reviews are bad for companies that provide quality service and products. People do not use their products, but positive comments about spam are not good for users because they cheat and undermine trust. As a result, researchers have discovered different ways to identify misleading comments. Some of them are very effective and easy to use. But some people have to do more work to improve their performance. Different researchers use different sets of data, such as some using Yelp data sets, some using Twitter data, and some using Amazon datasets. Different methods have been implemented on these data sets and some effective methods have been found. However, researchers still need more work to improve their effectiveness. As my

research question, 01 is to apply sentiment analysis using NLP to find the response of tweets in provides Positive, Negative, Strongly Positive, Weakly Positive, Natural, Weakly Negative, and Strongly Negative results. These results are on the tweeter based data analysis by taking the recent 500 tweets of Imran khan.

As we all know, the use of the Internet is increasing rapidly every day. People use the Internet to find anything and can check every detail. Nowadays, before starting any job, people must first browse, check their reputation, and comment on any work, and then decide to do so. As a result, spam comments have increased very quickly. These companies hire spammers to write spam comments about products designed to improve their business, but sometimes they hire spammers to disrupt the business of other competitors. There are two types of comments, one positive and the other negative. Different researchers have studied this and tried to find different effective methods to detect these false comments.

Because false reviews are bad for companies that provide quality service and products. People do not use their products, but positive comments about spam are not good for users because they cheat and undermine trust. As a result, researchers have discovered different ways to identify misleading comments. Some of them are very effective and easy to use. But some people have to do more work to improve their performance. Different researchers use different sets of data, such as some using Yelp data sets, some using Twitter data, and some using Amazon datasets. Different methods have been implemented on these data sets and some effective methods have been found. However, researchers still need more work to improve their effectiveness.

Data set polarities

Table 1: Polarities data set

Data set	Positive	Weekly positive	Strongly positive	Negative	Weekly negative	Strongly negative	Neutral
#ImranKhanPTI	15	22	4	11	18	0	112
#MubashirLuqman	23	43	13	7	23	6	207
# SabirShakir	39	50	12	20	31	4	342
# HamidMirPAK	40	95	17	22	37	7	279
#MoeedPirzada	38	105	13	23	60	9	245
#WalterCronkite	3	5	0	0	1	1	23
# LesterHoltNBC	53	100	31	26	46	12	231
# camanpour	50	72	18	29	22	59	257
# FWhitfield	22	21	13	7	10	1	81
#realBobWoodward	31	46	14	17	31	4	192

The above table show the polarities against the abstract or top-level keyword in another word primary level hash tag of 500 tweets of each journalist to search and find the polarities using NLP with the commonly used python library TextBlob in the next phase applying all results mapped on the graph for the final results.

Table 2: Find Polarities using NLP

Data set	Positive	Weekly positive	Strongly positive	Negative	Weekly negative	Strongly negative	Neutral
#ImranKhanPTI	3.00%	4.40%	0.80%	2.20%	3.60%	0.00%	22.40%
#MubashirLuqman	4.60%	8.60%	2.60%	1.40%	4.60%	1.20%	41.40%
# SabirShakir	7.80%	10.00%	2.40%	4.00%	6.20%	0.80%	68.40%
# HamidMirPAK	8.00%	19.00%	3.40%	4.40%	7.40%	1.40%	55.80%
#MoeedPirzada	7.60%	21.00%	2.60%	4.60%	12.00%	1.80%	49.00%
#WalterCronkite	0.60%	1.00%	0.00%	0.00%	0.20%	0.20%	4.60%
#LesterHoltNBC	10.60%	20.00%	6.20%	5.20%	9.20%	2.40%	46.20%
#camanpour	10.00%	14.40%	3.60%	5.80%	11.80%	2.40%	51.40%
#FWhitfield	4.40%	4.20%	2.60%	1.40%	2.00%	0.20%	16.20%
#realBobWoodward	9.20%	2.80%	3.40%	6.20%	0.80%	2.80%	38.40%

How people are reacting to journalist by analyzing 500 tweets on each. dataset of different Hashtags and Keywords from Twitter account in percentage. These hashtags are classified by their subjectivity. These are classified as positive, negative, and neutral on

a polarities basis. We calculate the polarities of the dataset and classify them accordingly. The table shows the result of each classification depending on the polarity of the data. The result shows most of the reviews are positive, some of them are negative and few reviews result as neutral.

In this way, I toke analysis on #MubasherLucman Hash tag and Mubasher Lucman keywords for sentiment analysis that provides me with all positive and negative aspects. So After analyzing #tags are most often used for elections or other events and happening. I analyzed the most commonly used keywords and most of the #tags and also on journalists like PM Imran khan as shown in the Graph chart along with line charts.

Imran Khan

This is a triple graphical analysis of the first 500 tweets of November 26, 2020.

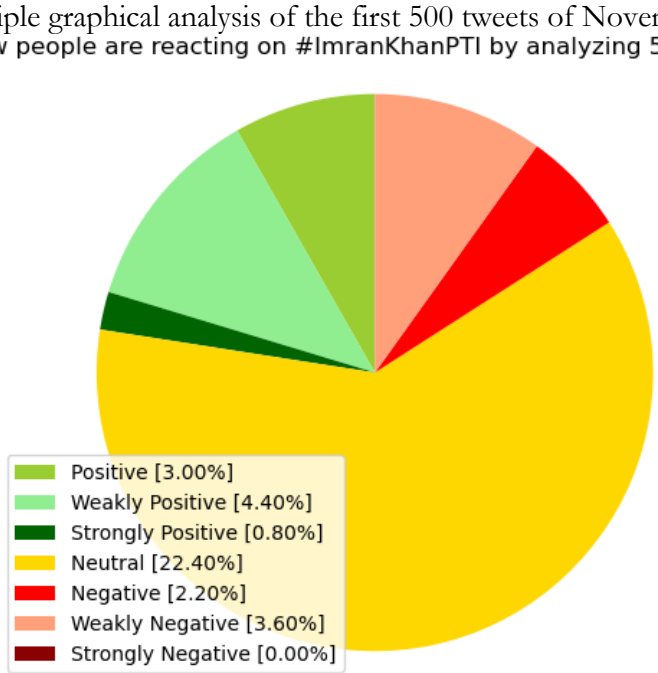

Figure 7: Imran khan Graph Chart Tweets Analysis

This is the result of the analysis applied to PM Imran Khan. This is a real-time sentiment analysis of Imran Khan using tweeter data.

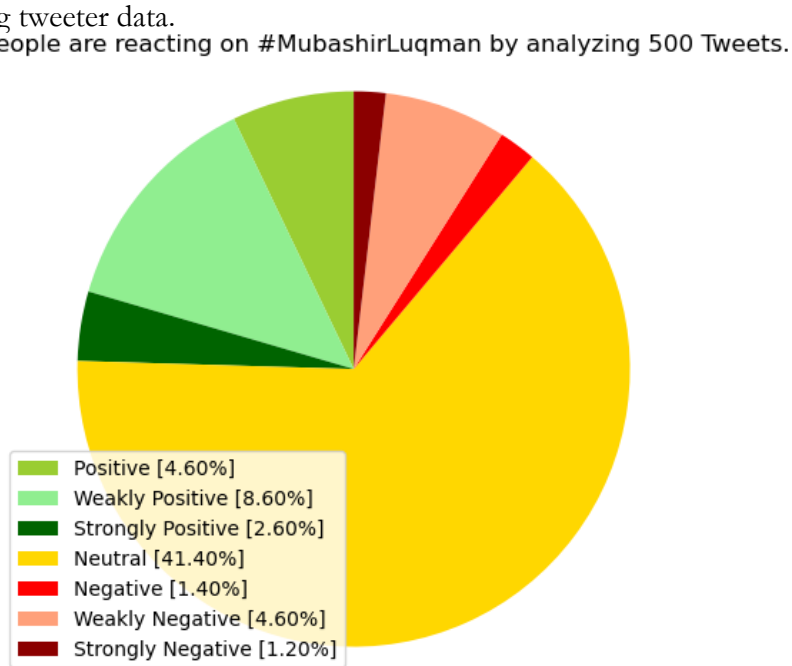

Figure 8: Mubashir Luqman Graph Chart Tweets Analysis

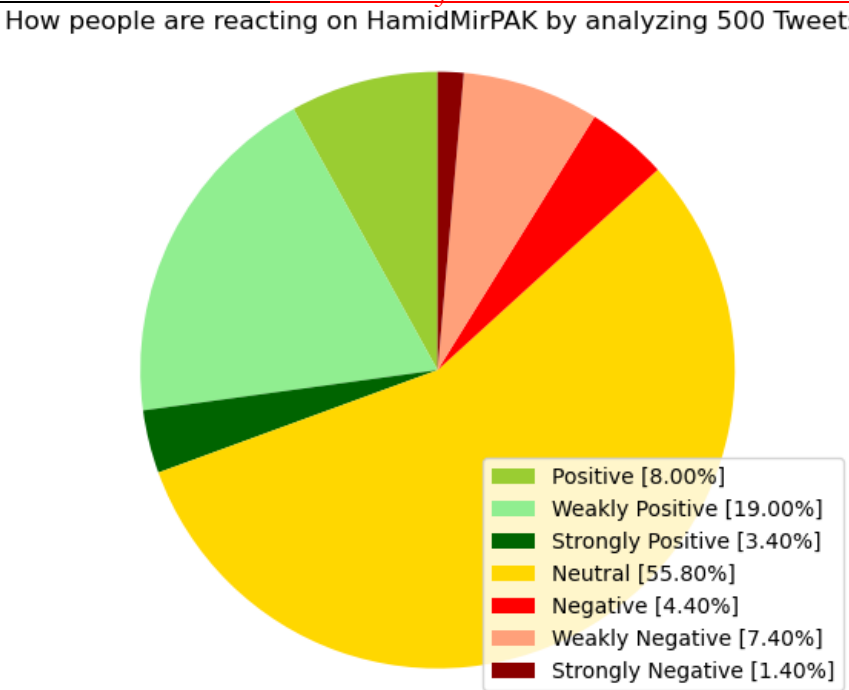

Figure 9: Hamid Mir Graph Chart Tweets Analysis

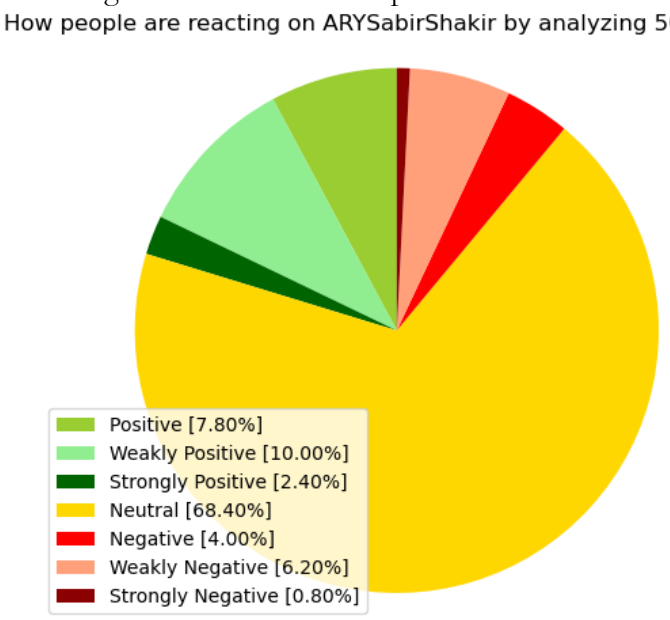

Figure 10: Sabir Shakir Graph Chart Tweets Analysis

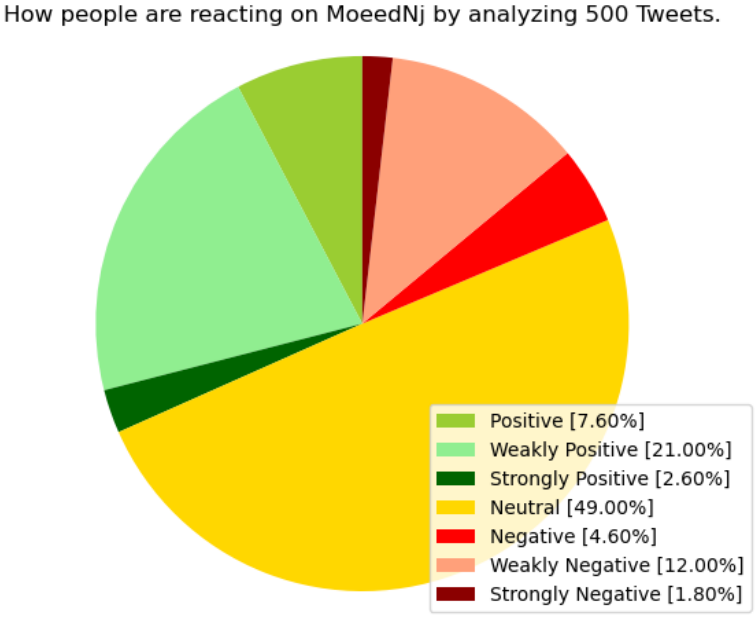

Figure 11: Mooed pir zadaGraph Chart Tweets Analysis

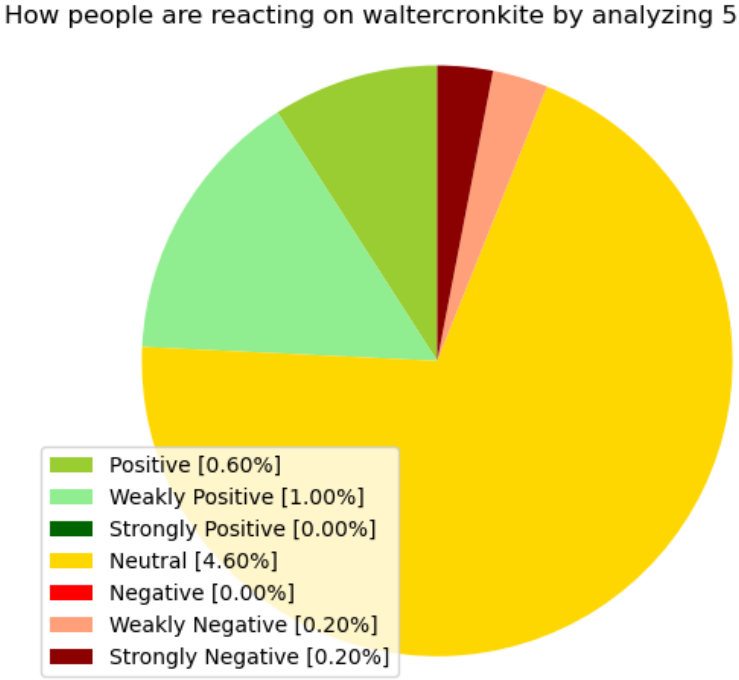

Figure 12: Walter Cronk Graph Chart Tweets Analysis

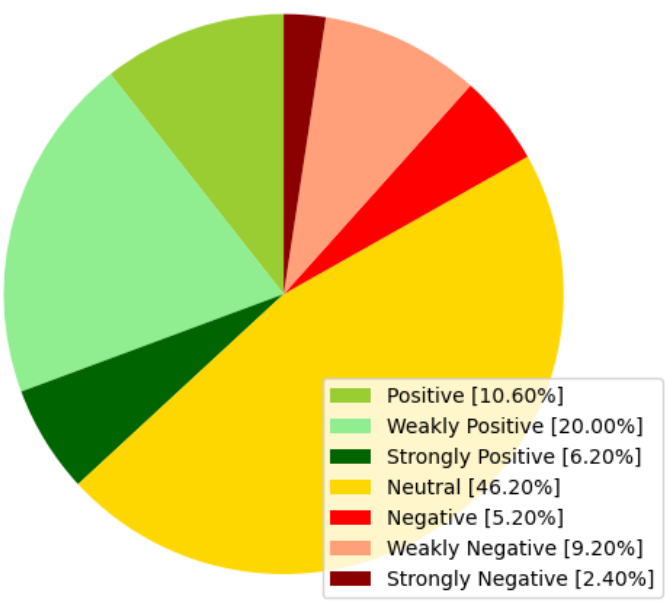

Figure 13: Lester Holt Graph Chart Tweets Analysis

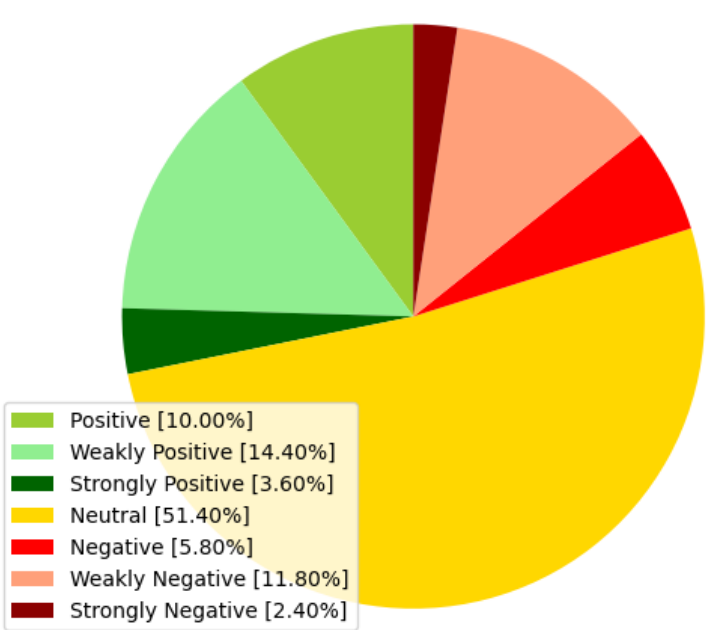

Figure 14: Camanpour Graph Chart Tweets Analysis

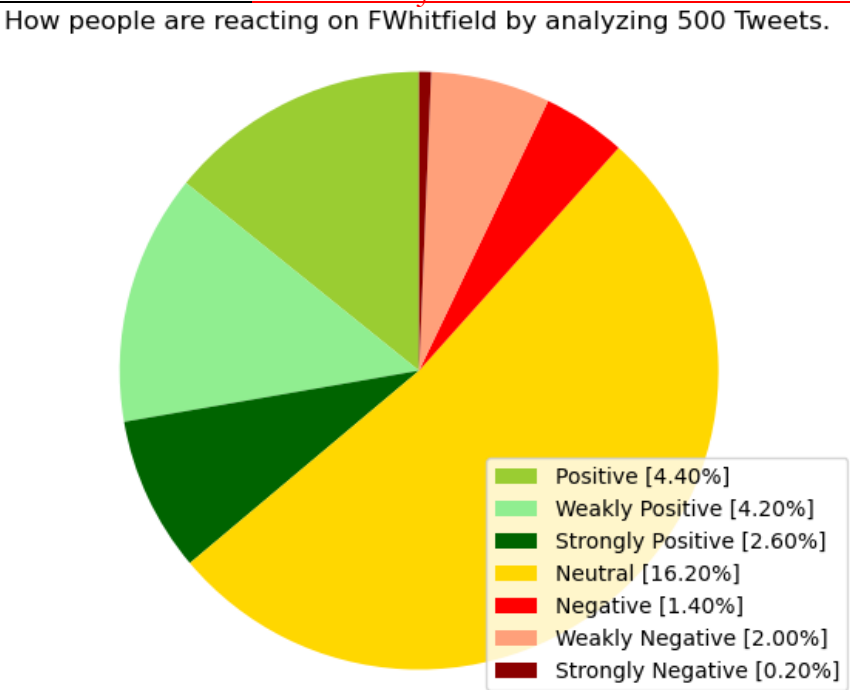

Figure 15: FWhitfield Graph Chart Tweets Analysis

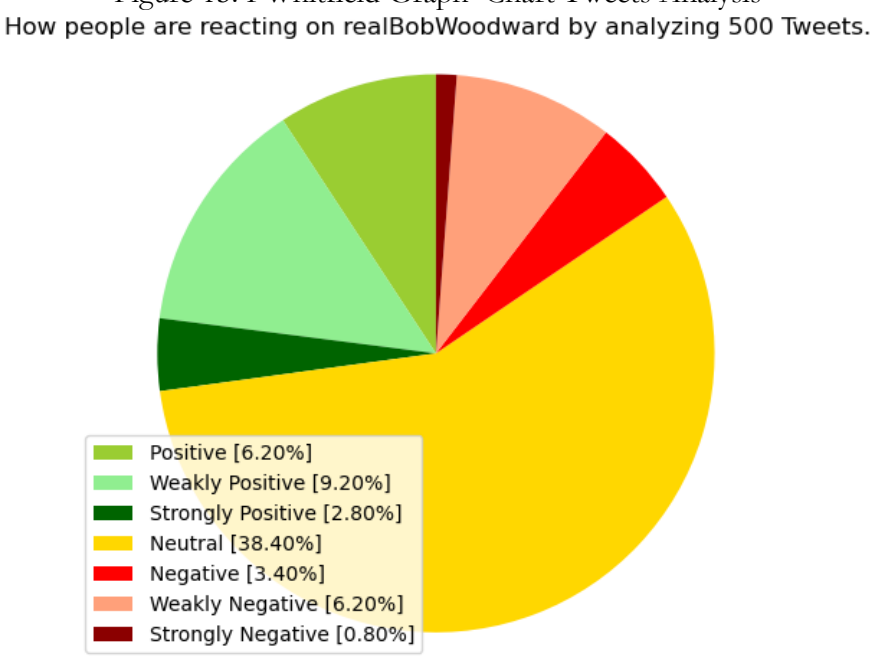

Figure 16: BobWoodward Graph Chart Tweets Analysis

It is a sentiment analysis of the tweets, transmissions, tastes, and number of accounts involved in the analysis. The results show that few accounts play a role in the negatives, more accounts are positive and some accounts show natural results.

CONCLUSION

Opinions give us suggestions of who our contacts are and who we believe. We also love to explain to people our opinions, as a way to define our traits. Sentiment analysis, occasionally also referred to as opinion drawing out, seeks to perform just this. Generate data-driven method to categorize the polarity of a text. Sentiment classifiers help to track the user options about a different topics like predicting the result predicting the emotions, and aggressive hate speech of journalists. Cross-language offers us a path to carry out sentiment analysis in an under-resourced language that does not contain any annotated data accessible. Resource-rich languages like English have high-feature annotated data for a lot of sentiment everyday jobs and domains. On the other hand, under-resourced languages also completely require annotated data or have just a few resources for exact domains or sentiment everyday jobs. This motivation require increasing sentiment analysis methods able to leverage data annotated in new languages. We developed an analysis between journalists and the public mostly, trending features, a passing a bundle of steps. We successfully evaluate the prediction on the base of the journalist's area using language sentiments and interest ratio calculated with the help of comments and tweet-like ratio. One of the main advantages of this thesis is taking the top-level decisions on the base of social media analysis.

References

[1] A. Hellweg, "Social Media Sites of Politicians Influence Their Perception by Constituents," J. Undergrad. Res. Commun., vol. 2, no. 1, pp. 22–36, 2011.

[2] A. Pandey, S. Mukherjee, S. Bajaj, and A. Jonnalagadda, Identification of meme genre and its social impact, vol. 105. Springer Singapore, 2019.

[3] L. Bing, K. C. C. Chan, and C. Ou, "Public sentiment analysis in twitter data for prediction of a company's stock price movements," Proc. - 11th IEEE Int. Conf. E-bus. Eng. ICEBE 2014 - Incl. 10th Work. Serv. Appl. Integr. Collab. SOAIC 2014 1st Work. E-Commerce Eng. ECE 2014, pp. 232–239, 2014, doi: 10.1109/ICEBE.2014.47.

[4] G. Enli and C. A. Simonsen, "'Social media logic' meets professional norms: Twitter hashtags usage by journalists and politicians," Inf. Commun. Soc., vol. 21, no. 8, pp. 1081–1096, 2018, doi: 10.1080/1369118X.2017.1301515.

[5] M. Collins and A. Karami, "Social Media Analysis For Organizations: Us Northeastern Public And State Libraries Case Study," pp. 1–6, 2018.

[6] M. Bilal, H. Israr, M. Shahid, and A. Khan, "Sentiment classification of Roman-Urdu opinions using Naïve Bayesian, Decision Tree and KNN classification techniques," J. King Saud Univ. - Comput. Inf. Sci., vol. 28, no. 3, pp. 330–344, 2016, doi: 10.1016/j.jksuci.2015.11.003.

[7] M. Knight, "Journalism as usual: The use of social media as a newsgathering tool in the coverage of the Iranian elections in 2009," J. Media Pract., vol. 13, no. 1, pp. 61–74, 2012, doi: 10.1386/jmpr.13.1.61_1.

[8] C. Bauckhage, "Insights into Internet Memes," Proc. Int. Conf. Weblogs Soc. Media, pp. 42–49, 2011.

[9] S. He et al., "Massive meme identification and popularity analysis in geopolitics," 2019 IEEE Int. Conf. Intell. Secur. Informatics, ISI 2019, no. 2017, pp. 116–121, 2019, doi: 10.1109/ISI.2019.8823294.

[10] K. Amarouche, H. Benbrahim, and I. Kassou, "Product Opinion Mining for Competitive Intelligence," Procedia Comput. Sci., vol. 73, no. Awict, pp. 358–365, 2015, doi: 10.1016/j.procs.2015.12.004.

[11] N. A. Morton and Q. Hu, "Implications of the fit between organizational structure and ERP: A structural contingency theory perspective," Int. J. Inf. Manage., vol. 28, no. 5, pp. 391–402, 2008, doi: 10.1016/j.ijinfomgt.2008.01.008.

[12] L. Shifman, "Memes in a digital world: Reconciling with a conceptual troublemaker," J. Comput. Commun., vol. 18, no. 3, pp. 362–377, 2013, doi: 10.1111/jcc4.12013.

[13] W. Fan and M. D. Gordon, "The power of social media analytics," Commun. ACM, vol. 57, no. 6, pp. 74–81, 2014, doi: 10.1145/2602574.

[14] X. Deng and R. Chen, "Sentiment analysis based online restaurants fake reviews hype detection," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 8710 LNCS, pp. 1–10, 2014, doi: 10.1007/978-3-319-11119-3_1.

[15] R. Boyle, "Sports Journalism: Changing journalism practice and digital media," Digit. Journal., vol. 5, no. 5, pp. 493–495, 2017, doi: 10.1080/21670811.2017.1281603.

[16] M. Crawford, T. M. Khoshgoftaar, and J. D. Prusa, "Reducing feature set explosion to facilitate real-world review spam detection," Proc. 29th Int. Florida Artif. Intell. Res. Soc. Conf. FLAIRS 2016, pp. 304–309, 2016.

[17] L. Xie, J. R. Kender, M. Hill, and J. R. Smith, "Xie et al - Visual memes is social media.pdf," pp. 53–62, 2011.

[18] N. Diakopoulos, M. Naaman, and F. Kivran-Swaine, "Diamonds in the rough: Social media visual analytics for journalistic inquiry," VAST 10 - IEEE Conf. Vis. Anal. Sci. Technol. 2010, Proc., pp. 115–122, 2010, doi: 10.1109/VAST.2010.5652922.

[19] P. B. Brandtzaeg, A. Følstad, and M. Á. Chaparro Domínguez, "How Journalists and Social Media Users Perceive Online Fact-Checking and Verification Services," Journal. Pract., vol. 12, no. 9, pp. 1109–1129, 2018, doi: 10.1080/17512786.2017.1363657.

[20] G. G. Chowdhury, "Natural language processing," Annu. Rev. Inf. Sci. Technol., vol. 37, pp. 51–89, 2003, doi: 10.1002/aris.1440370103.

[21] S. A. Eldridge, K. Hess, E. Tandoc, and O. Westlund, "Editorial: Digital Journalism (Studies)–Defining the Field," Digit. Journal., vol. 7, no. 3, pp. 315–319, 2019, doi: 10.1080/21670811.2019.1587308.

[22] M. Kim, L. Xie, and P. Christen, "Event diffusion patterns in social media," ICWSM 2012 - Proc. 6th Int. AAAI Conf. Weblogs Soc. Media, pp. 178–185, 2012.

[23] R. Barbado, O. Araque, and C. A. Iglesias, "A framework for fake review detection in online consumer electronics retailers," Inf. Process. Manag., vol. 56, no. 4, pp. 1234–1244, 2019, doi: 10.1016/j.ipm.2019.03.002.

[24] P. H. C. Guerra, D. Guedes, W. Meira, C. Hoepers, M. H. P. C. Chaves, and K. Steding-Jessen, "Spamming chains: A new way of understanding spammer behavior," 6th Conf. Email Anti-Spam, CEAS 2009, 2009.

Copyright © by authors and 50Sea. This work is licensed under Creative Commons Attribution 4.0 International License.