



A STATISTICAL APPROACH USING ML PERSPECTIVES FOR CLASSIFICATION AND MINING ON TWITTER DATASETS

KHAN NIKHAT KHANAM SARDAR¹, ASST.PROF. V. S. KARWANDE²

ME Student, Department of Computer Science & Engineering, EESGOI, India¹

HOD, Assistant Professor, Department of Computer Science and Engineering, EESGOI, India. ²

Abstract: With more than 300 million members, Twitter is a well-known online social networking platform where a tonne of information is shared daily. The most significant element of Twitter is its ability for users to tweet about events, situations, feelings, opinions, or even something new in real time. Since there is no framework in place to guarantee veracity or reliability, anyone can say whatever they want. This kind of Twitter war could be an easy way to attack your opponents and allow them to attack you in return. This survey covered a number of machine learning and hate speech identification apps that help identify sarcasm and shame tweets. With the rise of public shaming events and online social network expansion, it is necessary to evaluate shaming tweets, divide them into different categories, and restrict them in order to stop them.

Keywords: Naive Bayes(NB); Application programming interface(API); Online social networks (OSNs); K-nearest neighbours (KNN); Random forest(RF).

I INTRODUCTION

Because OSNs are poisonous, the victim's ignorance of the truth usually leads to embarrassment, financial loss, or both. The literature has extensively studied hate speech, bullying, profanity, flame, trolling, and other uninteresting speech in online social networks. However, from a computational perspective, public shaming—critiquing an individual for defying accepted social standards in an effort to make them feel ashamed—has not gotten much attention. Still, during the past few years, the number of these cases has been rising rapidly. Such occurrences are viral, as seen by the vast number of comments that are regularly used to shame a

victim who is essentially unknown. When American Internet Company public relations representative Justine Sacco tweeted, "Heading to Africa. My goal is to stay clear of AIDS. I'm kidding. When she shared her message, "I'm white!" she had just 170 followers. With a barrage of criticism, the incident became one of the most discussed topics on Twitter and the Internet in general within hours. She'd lost her job even before she got to South Africa. In his song "So You've Been Publicly Shamed," Jon Ronson recounts the experiences of numerous individuals who have been subjected to online bullying. The author's opinion is supported by the numerous examples of victims receiving sentences that are out of proportion to the seriousness of the crimes they committed. For every incidence they've examined, they've also put together a list of victims, the year it happened, the action that led to public humiliation, the medium that set it off, and the immediate consequences of the incident. The trigger is the first word or deed spoken by the victim who is responsible for the public humiliation. The "medium of triggering" was the first communication channel that the general public used to learn about the "Trigger." The impacts on the victim that happen during or shortly after the incident are listed under "Immediate repercussions." From now on, the embarrassing incident in question shall be referred to by the two-letter acronyms of the victim's name. The author offered a method for automating the process of identifying shamers in tweets from the victim's point of view, emphasizing two crucial elements: events and shamers. Additionally, there are six sorts of shameful tweets: abusive, comparison, passing judgment, religious/ethnic, sarcastic/joking, and whataboutery. Each tweet is classified as either non-shaming or belongs to one of these categories.

It is discovered that most people who remark on a humiliating occurrence have a tendency to shame the victim. Shamers will also see a quicker increase in count than non-shamers.

II LITERATURE SURVEY

In this research, Sarcasm is a distinct assessment type in this study that is made up of phrases that typically convey the opposite of what you really want to express. Sarcasm may be used to be clever, insult, or anger people in a number of contexts. Sometimes, sarcasm is employed inside a message to convey information that is comprehended. Sarcasm may be used in many different contexts, much like analysis or humor. Whatever the case, people find it quite challenging to recognize sarcasm in any given circumstance. Consequently, sarcasm detection improves understanding of the user's sentiment analysis, which is predicated on information gathered from social media platforms such as Facebook and Twitter. Research interest in automated sarcasm identification has increased due to the advantages of sarcasm detection for opinion mining. In automated sarcasm recognition, the program looks for neutral or sarcastic language. The study paper is divided into two sections. In the first step, it collects features related to sentiments and punctuation; in the second part, the chi-square test is used to determine which characteristics are most valuable. The second step takes 200 top tf-idf features and combines them with punctuation-related traits and emotions to find sarcastic material inside a tweet. In the first approach, the voting classifier obtains the maximum accuracy of 83.53 percent, while the Support Vector Machine technique earns the best accuracy of 74.59 percent. [1].

In this research, While social media data, such as that from Twitter, is valuable for research, its replicability makes it challenging to build compact, self-contained datasets that are vast in size. Data recording is a common problem in social media-based research that should be standardized. Unfortunately, this is not often done. This paper presents the results of a long-term evaluation study that captured the whole public sample of the German and English Twitter streams. It suggests a method of recording that chunks a linear stream of events to minimize duplication. If events are seen many times inside a chunk's time span, just the most recent observation is applied to the chunk. A 10-gigabyte raw Twitter dataset including 1.2 million Tweets from 120,000 users, gathered between June and September 2017, was used to estimate expected compression rates. It was shown that all of the events, information, and connections between individual events could be retained in the created datasets with just 10 to 20 percent of the original data size. The standardized and repeatable curation of large-scale (even national) self-contained and tiny datasets of social networks

for research is made possible by this kind of redundancy reduction recording. [2].

In this research,, One hot subject at the moment is the difficulty of recognizing and categorizing tweets linked to extremism. Extremist gangs have a history of recruiting new members and disseminating their ideologies using social media sites like Facebook and Twitter. This project aims to provide a framework for the analysis of information on terrorism, with a focus on categorizing tweets into extremist and non-extremist groups. Our method classifies tweets as either extremist or non-extremist by utilizing deep learning-based sentiment analysis techniques on user-generated social media posts on Twitter. The experiment's results are encouraging and provide a foundation for further research. [3].

In this research, Depression has long been a cause for concern in our society and continues to be a contentious topic among academics worldwide. Predicting depressed moods is still difficult, despite a wealth of study on the subject of understanding individual moods, including sadness, anxiety, and stress, as well as activity logs gathered by ubiquitous computing devices like smartphones. In order to predict suicide behaviors based on depression levels, we provide in this research a depression analysis and suicidal ideation detection approach. The PHQ-9 (Parent Health Questionnaire)-style questionnaires, which included questions like "What is your age?" were used to collect real-time data from parents and students. Are you a regular student or college attendee? and transformed it into pertinent information including age, sex, attendance at school, and other details. Next, classification machine algorithms are used to train and classify depression into five phases based on the severity of the condition. There are four main levels of severity: mild, moderate, fairly severe, and severe. With an accuracy of 83.87 percent, the XGBoost classifier produced the best results in this dataset. Additionally, information was gathered through the collection of tweets, which were then analyzed using classification algorithms to ascertain whether or not the tweeter was sad. The Logistic Regression classifier yielded the highest accuracy, 86.45 percent. [4].

In this research, People that use sarcasm communicate in a subtle way by saying the exact opposite of what is meant to be said. Prior studies on sarcasm recognition were conducted on text. Nonetheless, an increasing number of social media platforms, like Twitter, enable users to send multimodal messages, which include text, images, and videos. It is insufficient to use texts alone to detect sarcasm in multimodal communication. In this paper, the authors concentrate on multimodal sarcasm detection for tweets that have both text and images. The authors suggest a multi-modal hierarchical fusion approach to address this issue by treating text input, picture characteristics, and picture attributes as three modalities. Our model uses a bidirectional LSTM

network and attribute features to extract text features after first extracting picture and attribute features. Three modalities' features are rebuilt and combined into one feature vector for prediction. We create a multi-modal sarcasm detection dataset using Twitter. The results of the dataset evaluation demonstrate the usefulness of the three modalities and the efficacy of our proposed model. [6].

In this research, Social networking services such as Facebook, Instagram, MySpace, Twitter, and Facebook are becoming increasingly effective tools for communication. These have become a part of our everyday existence. Individuals may feel more a part of their society when they communicate their opinions and behaviors to their social circle. But there are drawbacks to this freedom of speech as well. Aggressive behavior is occasionally displayed on social media, which hurts the sentiments of the intended victims. Some forms of cyberbullying are driven by physical, racial, or sexual limitations. Sufficient surveillance is therefore necessary to handle such situations. As a microblogging site, Twitter frequently experiences cyberbullying. Conversely, tweets are raw language that frequently include misspelled and restricted phrases. This work presents the Bag-of-Phonetic-Codes model as a unique cyberbullying detection method. Word pronunciation can be used as a feature to identify forbidden terms and fix misspelled words. If duplicate words are correctly identified, the lexicon of terms and hence the feature space may be less. This proposed work was inspired by the well-known Bag-of-Words approach for extracting textual features. Phonetic codes were produced by use of the Soundex Algorithm. In addition to the suggested model, experiments with supervised and unsupervised machine learning algorithms were carried out on a number of datasets in order to gain a deeper understanding of the methods and problems related to cyberbullying detection. [8].

In this research, Twitter is essential for distributing information and influencing people's opinions. Twitter is an easy-to-use platform that enables people to share their opinions and thoughts. Studies have demonstrated that monitoring and evaluating social media sentiment can help forecast certain political developments. As a result, politicians have been enticed to solicit public opinion on their proposals. Sentiment analysis of political tweets becomes increasingly significant as a result. Twitter users spoke about US Republican presidential candidate Donald Trump after his primary debates. In order to determine if tweets supporting or opposing Trump are being sent, it also examines how people are feeling about him after his debates. [10].

In this research, A sophisticated kind of irony that is common on social media and microblogging sites is sarcasm. It is most frequently used to convey implicit information that is sent by a person inside a discussion. Sarcasm may be employed for a number of purposes, such as ridiculing and

criticism. But even humans have trouble identifying it. Therefore, sarcastic comment detection can improve automatic sentiment analysis of data collected from social media or microblogging websites. Sentiment analysis is the process of identifying and compiling opinions and attitudes that Internet users have expressed about a certain topic. In this paper, introduce a pattern-based method for sarcasm detection on Twitter. Provide four sets of traits that symbolize the many types of sarcasm that have been found. To discern between sarcastic and non-sarcastic tweets, use these. Our proposed approach achieves an accuracy of 83.1 percent with a precision of 91.1 percent. also examine each suggested feature set's importance and evaluate how much more valuable it adds to the classification. Emphasize the value of pattern-based features in particular for sarcastic speech detection. [14].

In this research, Surveys have long been used to gather information on the inclination toward mental health illnesses, as well as the ebbs and flows of feelings, ideas, and behaviors. Nevertheless, surveys are unable to provide current information about people's or groups' emotional states. Large data sets, like WHO statistics, are only collected once a year; in contrast, real-time analysis of reported mood is possible on social networking sites like Twitter. These trends may be used by the mental health research community to determine where and when there is the greatest need and demand. The authors talk about the validation results against recognized patterns of mood shift, as well as the "authors Feel" approach for analyzing changes in emotional expression on a global and regional scale. Using principal component analysis (PCA), the results showed a dominant in-phase pattern for all emotions, which was modified by anti-phase patterns for "positive" and "negative" emotions. The first three primary components explained more than 90% of the variation in the data. The primary weekly and daily oscillations were removed using PCA, which also allowed for the identification of significant events within the data, with z-scores indicating emotional expressiveness exceeding 80 standard deviations from the mean. The burden of anxiety and suicide rates appeared to be associated to emotional expression, but no connections were observed when the authors matched emotional expression to national WHO figures. [15].

III. SYSTEMS ARCHITECTURE

Our aim is to automatically categorize tweets into the six categories mentioned below using the system model architecture. Figure 1 shows the main functional components involved in the automated classification of tweets that shame people. For every category, the labeled training set and test set of tweets undergo the preprocessing and feature extraction procedures. The training data is used to train the Random Forest (RF), Naive Bayes (NB), and K-nearest

neighbors (KNN) classifiers. The trained RF's precision scores are analyzed on the test set. Based on these scores, the classifiers are arranged hierarchically. After preprocessing and feature extraction, a new tweet is submitted to the trained classifiers, labeled with the class of the first classifier to classify it as positive. A tweet is deemed nonshameful if it receives a negative classification from every classifier.

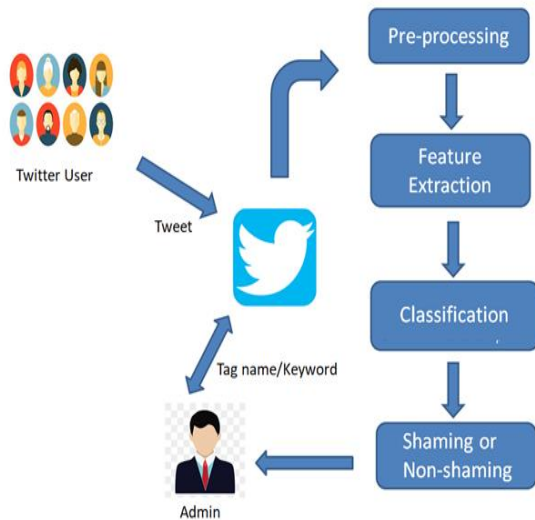


Figure No 3.1: System Architecture

IV EXPERIMENTAL RESULTS

Massive amounts of tweets from various embarrassing incidents spread across years were gathered using the Twitter 1 percent stream, a Twitter search application programming interface (API). The objective assigned to the annotators was to classify the tweets into one of the six degrading categories or mark them as NS. Information on the humiliating incidents that were collected; the term "#Annotated" denotes the quantity of tweets that were manually labeled for every occurrence. No annotated data is available for events LD, MT, or PC. "#unique tweets" is the total number of unique tweets that have been collected for an event. We do not expressly include retweets in the data set because they are given the original tweet's title.

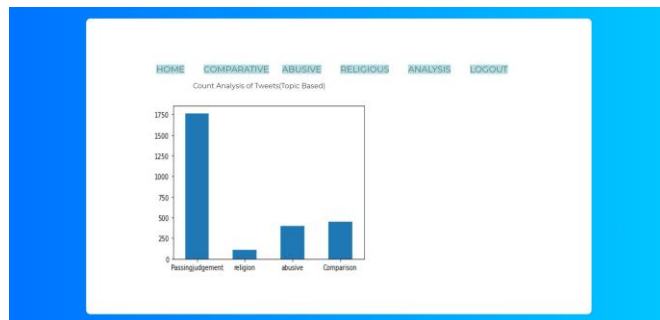


Figure No 4.1: No. of Shammer & Non-Shammer Comparison.

V CONCLUSION

All things considered, the need for the application is increasing with the number of social media sites that are online, the frequency of public humiliation incidents, and the volume of criticism directed against the callousness of website owners. In order to identify shammers and shaming tweets more easily, this study looks at several uses of machine learning and hate speech detection. A new approach that may be offered once the literature has been reviewed is able to classify offensive remarks into different categories and ban certain tweets if they are deemed offensive.

REFERENCES

- 1.Shakeel Ahmad,Muhammad Zubair Asghar,Fahad M. Alotaibi and Irfanullah Awan,"Detection and classification of social media-based extremist affiliations using sentiment analysis techniques",Springer Open,2019.
- 2.Swati Jain, Suraj Prakash Narayan, Rupesh Kumar Dewang, Utkarsh Bhartiya, Nalini Meena and Varun Kumar,"A Machine Learning based Depression Analysis and Suicidal Ideation Detection System using Questionnaires and Twitter",IEEE,2019.
- 3.Liyuan Liu,Jennifer Lewis Priestley,Yiyun Zhou,Herman E. Ray and Meng Han,"A2Text-Net: A Novel Deep Neural Network for Sarcasm Detection", IEEE First International Conference on Cognitive Machine Intelligence (CogMI),2019.
- 4.Yitao Cai,Huiyu Cai and Xiaojun Wan,"Multi-Modal Sarcasm Detection in Twitter with Hierarchical Fusion Model",57th Annual Meeting of the Association for Computational Linguistics,August 2,2019.
- 5.Muhammad Abulaish and Ashraf Kamal,"Self-Deprecating Sarcasm Detection: An Amalgamation of Rule-Based and Machine Learning

- Approach", IEEE/WIC/ACM International Conference on Web Intelligence (WI), 2018.
6. Ankita Shekhar and M Venkatesan, "A Bag-of-Phonetic-Codes Model for Cyber Bullying Detection in Twitter", IEEE International Conference on Current Trends toward Converging Technologies, 2018.
7. Saurabh Porwal, Gaurav Ostwal, Anagha Phadtare, Manisha V. Marathe and Mohini Pandey, "Sarcasm Detection using Recurrent Neural Network", IEEE Second International Conference on Intelligent Computing and Control Systems (ICICCS), 2018.
8. Malak Abdullah and Mirsad Hadzikadic, "Sentiment Analysis of Twitter Data: Emotions Revealed Regarding Donald Trump during the 2015-16 Primary Debates", International Conference on Tools with Artificial Intelligence, IEEE, 2017.
9. Setra Genyang Wicana, Taha Yasin İbisoglu and Uraz Yavanoglu, "A Review on Sarcasm Detection from Machine-Learning Perspective", IEEE 11th International Conference on Semantic Computing, 2017.
10. Isobelle Clarke and Jack Grieve, "Dimensions of Abusive Language on Twitter", First Workshop on Abusive Language Online, pages 1–10, August 4, 2017.
11. Isaac David, Oscar S. Siordia and Daniela Moctezuma, "Features combination for the detection of malicious Twitter accounts", IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC), 2016.
12. Rahul Gupta, Jitendra Kumar, Harsh Agrawal and Kunal, "A Statistical Approach for Sarcasm Detection Using Twitter Data", IEEE International Conference on Intelligent Computing and Control Systems (ICICCS), 2020.
13. Nane Kratzke, "Redundancy Reduction in Twitter Event Streams", Preprints, 13 February 2020.
14. Mondher Bouazizi And Tomoaki Otsuki, "A Pattern-Based Approach for Sarcasm Detection on Twitter", IEEE Access, August 24, 2016.
15. Mark E. Larsen, Tjeerd W. Boonstra, Philip J. Batterham, Bridianne O'Dea, Cecile Paris, and Helen Christensen, "We Feel: Mapping emotion on Twitter", IEEE Journal of Biomedical and Health Informatics, 2015.
16. Neethu M S and Rajasree R, "Sentiment Analysis in Twitter using Machine Learning Techniques", 4th ICCNT, IEEE, 2013.
17. Dr. Tariq Mahmood, Tasmiyah Iqbal, Farnaz Amin, Wajeeta Lohanna and Atika Mustafa, "Mining Twitter Big Data to Predict 2013 Pakistan Election Winner", IEEE, 2013.