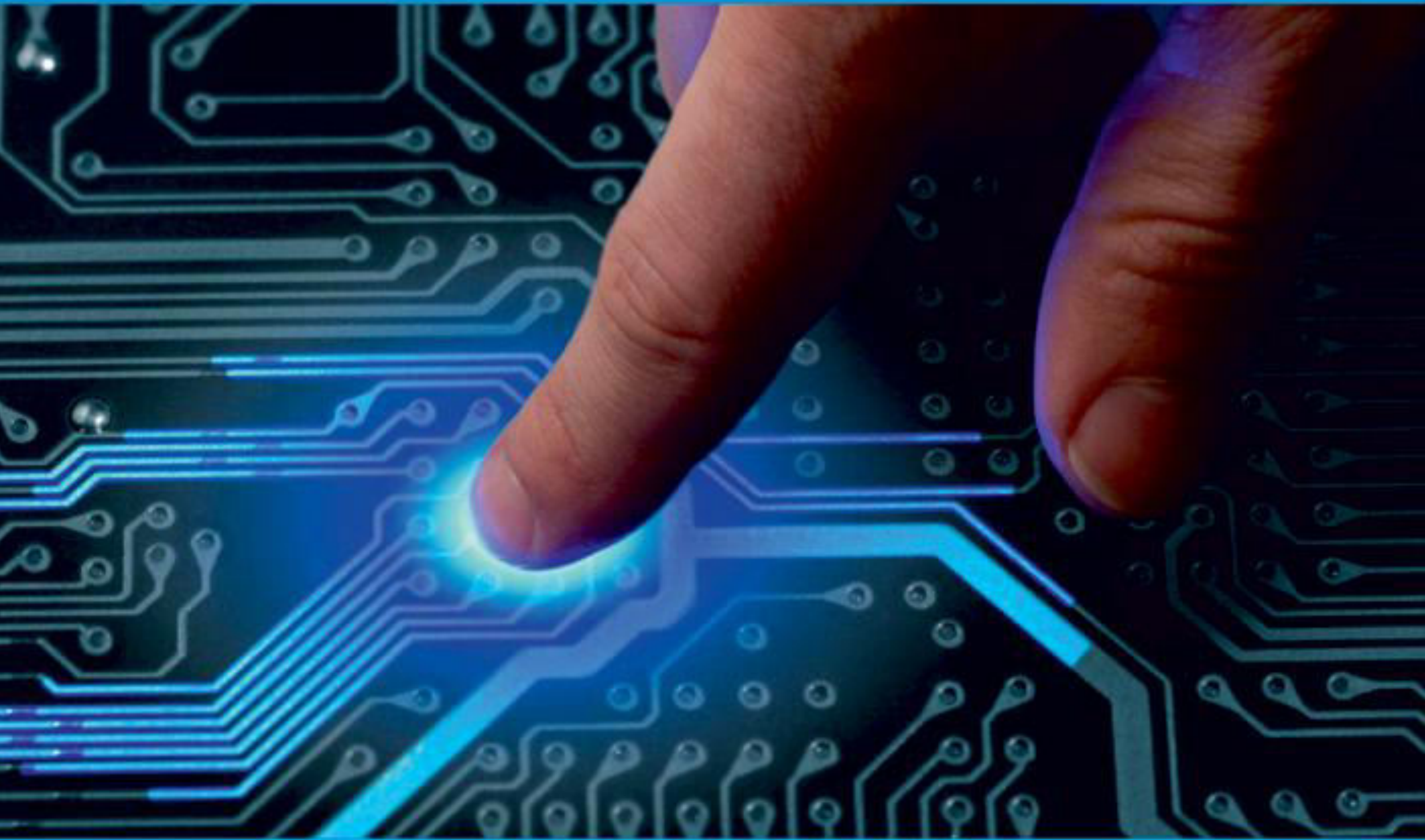
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 9, Issue 7, July 2021

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.542

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 9, Issue 7, July 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0907187 |

Cyber Bullying Detection using Machine Learning

Ms. Indumathi T, Ms. Keerthana N, Mrs. Bharathi Sri

UG Student, Dept. of CSE,Velammal Engineering College, Chennai, Tamil Nadu, India.

UG Student, Dept. of CSE,Velammal Engineering College, Chennai, Tamil Nadu, India.

Assistant Professor, Dept. of CSE,Velammal Engineering College, Chennai, Tamil Nadu, India.

ABSTRACT: The expansion of enthusiasm for utilizing online networking as a hotspot for investigate has inspired handling the test of naturally geo-locating tweets, given the absence of express area data in the lion's share of tweets. As opposed to much past work that has concentrated on area arrangement of tweets confined to a particular nation, here we attempt the errand in a more extensive setting by arranging worldwide tweets at the nation level, which is so far unexplored in a constant situation. We break down the degree to which a tweet's nation of starting point can be controlled by making utilization of eight tweet-inborn elements for order. Besides, we utilize two datasets, gathered a year separated from each other, to break down the degree to which a model prepared from recorded tweets can in any case be utilized for characterization of new tweets. With grouping probes each of the 217 nations in our datasets, and in addition on the best 25 nations, we offer a few bits of knowledge into the best utilization of tweet-intrinsic elements for an exact nation level grouping of tweets. We find that the utilization of a solitary component, for example, the utilization of tweet content alone the most broadly utilized element in past work clears out much to be wanted. Picking a proper blend of both tweet substance and metadata can really prompt considerable enhancements of in the vicinity of 20% and half. We watch that tweet content, the client's self-detailed area and the client's genuine name, all of which are characteristic in a tweet and accessible in a constant situation, are especially helpful to decide the nation of starting point.

KEYWORDS*:* cyber bullying, meta data, classification of tweets.

I. INTRODUCTION

In computer science, Machine learning is the study of computer algorithms that improve automatically through experience and by the use of data. This algorithm build a model based on sample data known as “training data”, in order to make predictions or decisions without being explicity programmed to do so. Extracted features with applied with Naïve Bayes algorithm, sentiment analysis algorithm, Ahocorasick algorithm etc..

Data plays a key role in any use case. 60% of the work of a data scientist lies in collecting the data. For beginners to experiment with machine learning, they can easily find data from Kaggle, UCI ML Repository, etc.To implement real case scenarios, you need to collect the data through web-scraping or (through APIs like twitter) or for solving business problems you need to attain data from clients (here ML engineers need to coordinate with domain experts to collect the data).Once the data is collected, we need to structure the data and store it in the database. This requires knowledge of Big data (or data engineer) which plays a major role here.

Once the data is collected you need to validate if the quantity is sufficient for the use case (if it is a time-series data, we need a minimum of 3–5 years of data).The two important things we do while doing a machine learning project are selecting a learning algorithm **and** training the model using some of the acquired data. So as humans, we naturally tend to make mistakes and as a result, things may go wrong. Here, the mistakes could be opting for the wrong model or selecting data which is bad. Now, what do I mean by bad data? Let us try to understand.Imagine your machine learning model is a baby, and you plan on teaching the baby to distinguish between a cat and a dog. So we begin with pointing at a cat and saying ‘ it’s a CAT**’** and do the same thing with a DOG (possibly repeating this procedure many times). Now the child will able to distinguish between dog and cat, by identifying shapes, colors, or any other features. And just like that, the baby becomes a genius (in distinguishing)!Similarly, we train the model with a lot of data. A child may distinguish the animal with less number of samples, but a machine learning model requires thousands of examples for even simple problems. For complex problems like Image Classification and Speech Recognition, it may require data in a count of millions.Therefore, one thing is clear. We need to train a model with sufficient Data.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 9, Issue 7, July 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0907187 |

The training data should be representative of the new cases to generalize well i.e., the data we use for training should cover all the cases that occurred and that is going to occur. By using a non-representative training set, the trained model is not likely to make accurate predictions. Systems which are developed to make predictions for generalized cases in business problem view are said to be good machine learning models. It will help the model to perform well even for the data which the data model has never seen .If the number of training samples is low, we have sampling noise which is unrepresentative data, again countless training tests bring sampling bias if the strategy utilized for training is defective.

II. RELATED WORKS

[1]Twitter, for instance, is one of the most popular social media platforms that allows its users to create their own profiles, upload their photos and videos, and send messages (both private and public). It has a wide reach, as any comments or posts can reach thousands of people, especially through "liking" and "sharing" mechanisms, and thus allowing cyber bullies to distribute nasty or unwanted information about their victims easily. Instagram allows its users to share photos and videos, to 'follow' others and support Stories.[2]The use of information and communication technologies, particularly social media has

revolutionized the manner in which people communicate and form relationships with one another, with statistics around the world indicating a high prevalence rate of social media applications This unfortunately, provides an avenue for anti-social behaviors such as misogyny.[3]Cyber bullying victimization, particularly involving young people has received an increasing level of scrutiny. For instance, Ask.fm (a platform that allows one to ask each other questions anonymously) had to launch new safety efforts when teenagers were bullied on their platform, resulting in several suicides. Likewise, Instagram recently introduced shadow banning online abusers (i.e., restricting a bully from posting or commenting on a post).[4]Facebook, for instance, is one of the most popular social media platforms that allows its users to create their own profiles, upload their photos and videos, and send messages (both private and public). It has a wide reach, as any comments or posts can reach thousands of people, especially through "liking" and "sharing" mechanisms, and thus allowing cyber bullies to distribute nasty or unwanted information about their victim easily.

III. METHODOLOGY

Machine learning is the study of computer algorithms that improve automatically through experience and by the use of data. This algorithm build a model based on sample data known as "training data", in order to make predictions or decisions without being explicity programmed to do so. Extracted features with applied with Naïve Bayes algorithm,sentiment analysis algorithm, Ahocorasick algorithm etc. Detecting social media bullying is done by John Hani et al. . They have used Neural Networks and classification models to detect and prevent social media bullying. After doing some research, they finally used NB and SVM for the detection of cyberbullying. For the proposed model they collected the dataset from the Kaggle. The proposed model is divided into 3 major steps:

1) Data Preprocessing

2) Feature Extraction

3) Classification

Preprocessing Steps:

o Tokenization

o Lowering Text

o Stop words and encoding cleaning

o Word correction

Feature Extraction: For feature extraction sentiment

analysis and TFIDF algorithms are used.

Classification: For classification SVM (Support Vector Machine) and NN classifiers are used.

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 9, Issue 7, July 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0907187 |

IV. MODULES

MODULE 1- GEOLOCATION BASED TWEET

In this modules, specific applications of a real-time, country-level tweet geo-location system. Our methodology enables us to perform a thorough analysis of tweet geo-location, revealing insights into the best approaches for an accurate country-level location classifier for geo-location based tweets.

MODULE 2- TWEET ANALYSIS

The user analyze the tweet's nature like positive negative or neutral by analyzing the containing words in the tweet and the each word we use in a regular basis comprises of words of positive and negative as well as neutral words. These words are initially inserted into the database for analysis

MODULE 3- SENTIWORD CALL

In this module the data to be analyzed for analysis is gets the score by calling the sentiword file because of the generation of score for individual words and this score is used for generation of analysis result. If any user post something malicious content in the post, that post has been blocked by the user. Admin first view the content posted by the user, based on the admin review post will be placed.

MODULE 4-ANALYSIS RESULT COMPARISON

The final score is generated by the result of the sentiword and this is useful in comparison of scores and this comparison results are kept some threshold to state the tweet as negative, positive or neutral one. Classification dependent upon the words posted on the twitter.

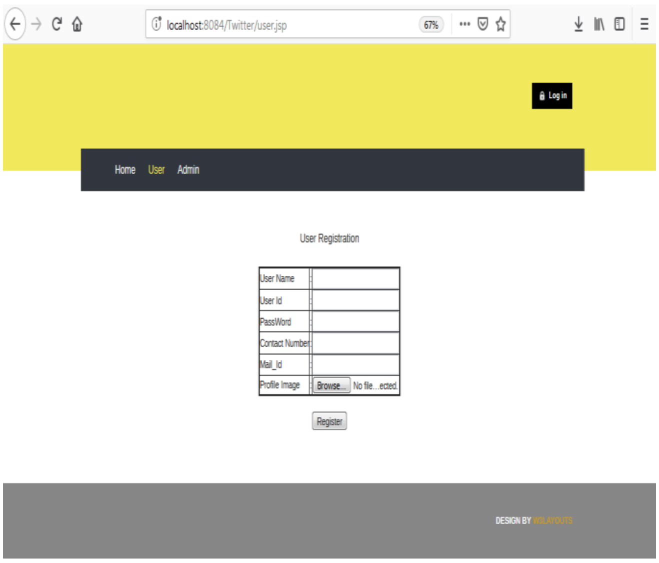


Fig 1:User Registration

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 9, Issue 7, July 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0907187 |

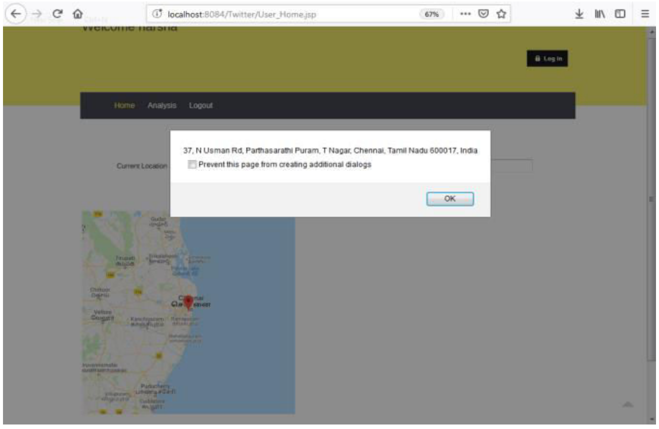


Fig 2:Location Based Tweets

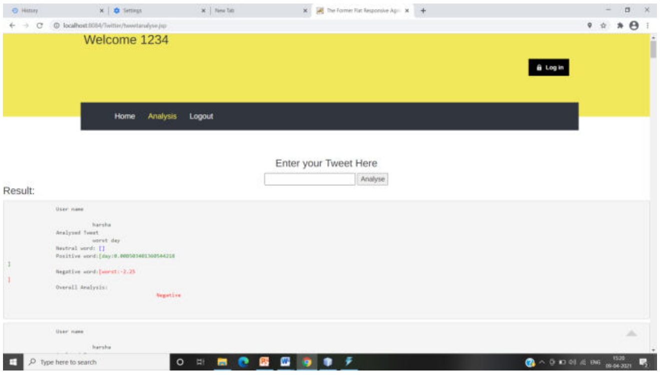


Fig 3: Tweet Analysis

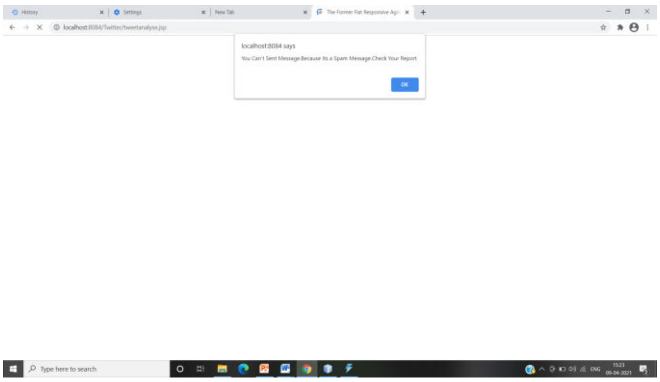


Fig 4: Spam Message

V. CONCLUSION

Mining sentiment polarities expressed in Twitter messages is a meaningful while challenging task. Most of the existing solutions to twitter sentiment analysis only consider textual information of Twitter messages, and cannot achieve satisfactory performance due to unique characteristics of Twitter messages. Although recent studies have shown that sentiment diffusion patterns have close relationships with sentiment polarities of Twitter messages, existing approaches basically only focus on textual information of Twitter messages, but ignore sentiment diffusion information. Inspired by recent work on fusion of knowledge from multiple domains, we take a first step towards combining textual and sentiment diffusion information to achieve better performance of Twitter sentiment analysis. To this end, we first analyze

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 9, Issue 7, July 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0907187 |

sentiment diffusion on Twitter by investigating a phenomenon called sentiment reversal, and find some interesting properties of sentiment reversals based on repost cascade trees and repost diffusion networks. We then build a sentiment reversal prediction model, and design a novel Twitter sentiment classification algorithm called SentiDiff. In SentiDiff, the inter-relationships between textual information of Twitter messages and sentiment diffusion patterns are considered, and the textual information based sentiment classifier and the sentiment reversal prediction model are combined in a supervised learning framework. The experiments on real-world dataset demonstrate that our proposed SentiDiff algorithm can help state-of-the-art textual information based sentiment analysis algorithms.

VI. FUTURE WORK

In the future study, we plan to analyze how sentiment diffusion patterns differ in different topics, and consider the topic information of Twitter messages when fusing textual and sentiment diffusion information.

REFERENCES

[1] O. Ajao, J. Hong, and W. Liu. A survey of location inference techniques on twitter. Journal of Information Science, 1:1–10, 2015.

[2] E. Amig´ o, J. C. De Albornoz, I. Chugur, A. Corujo, J. Gonzalo, T. Mart´ın, E. Meij, M. De Rijke, and D. Spina. Overview of replab 2013: Evaluating online reputation monitoring systems. In Proceedings of CLEF, pages 333–352. Springer, 2013.

[3] F. Atefeh and W. Khreich. A survey of techniques for event detection in twitter. Computational Intelligence, 31(1):132–164, 2015.

[4] H. Bo, P. Cook, and T. Baldwin. Geolocation prediction in social media data by finding location indicative words. In Proceedings of COLING, pages 1045–1062, 2012.

[5] J. Bollen, H. Mao, and A. Pepe. Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena. In Proceedings of ICWSM, pages 450–453, 2011.

[6] J. D. Burger, J. Henderson, G. Kim, and G. Zarrella. Discriminating gender on twitter. In Proceedings of EMNLP, pages 1301–1309, 2011.

[7] H.-w. Chang, D. Lee, M. Eltaher, and J. Lee. @ phillies tweeting from philly? predicting twitter user locations with spatial word usage. In Proceedings of ASONAM, pages 111–118, 2012.

[8] Y. Chen, J. Zhao, X. Hu, X. Zhang, Z. Li, and T.-S. Chua. From interest to function: Location estimation in social media. In Proceedings of AAAI, pages 180–186, 2013.

[9] Z. Cheng, J. Caverlee, and K. Lee. You are where you tweet: a content-based approach to geo-locating twitter users. In Proceedings of CIKM, pages 759–768, 2010.

10] R. Compton, D. Jurgens, and D. Allen. Geotagging one hundred million twitter accounts with total variation minimization. In IEEE Big Data, pages 393–401, 2014

[11] M. Conover, J. Ratkiewicz, M. R. Francisco, B. Gonc¸alves, F. Menczer, and A. Flammini. Political polarization on twitter. In Proceedings of ICWSM, pages 89–96, 2011.

[12] M. D. Conover, B. Gonc¸alves, J. Ratkiewicz, A. Flammini, and F. Menczer. Predicting the political alignment of twitter users. In IEEE PASSAT/SocialCom, pages 192–199, 2011.

[13] D. Doran, S. Gokhale, and A. Dagnino. Accurate local estimation of geo-coordinates for social media posts. arXiv preprint arXiv:1410.4616, 2014.

[14] M. Dredze, M. Osborne, and P. Kambadur. Geolocation for twitter: Timing matters. In Proceedings of NAACL-HLT, pages 1064–1069, San Diego, California, 2016.

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING