Volume 8, No. 5, May-June 2017
ISSN No. 0976-5697
International Journal of Advanced Research in Computer Science
SURVEY REPORT
Available Online at www.ijarcs.info

Survey of Cyberbulling Detection on Social Media Big-Data

Noopur Tarwani
Dept. of Computer Science
UIT RGPV
Bhopal, India

Prof. Uday Chorasia
Dept. of Computer Science
UIT RGPV
Bhopal, India

Dr. Piyush Kumar Shukla
Dept. of Computer Science
UIT RGPV
Bhopal, India

Abstract: Opinion mining or sentiment analysis is considered as an important application of NLP (Natural Language Processing). Opinion mining is extracting the views that people express online. Those Websites which permits social interaction and collaboration can be considered as social media site, including networking sites such as Facebook, MySpace, and Twitter. Such sites offer today's youth a platform for amusement, entertainment, thrill, correspondence and communication with friends and furthermore have developed radically and exponentially as of late. This is the reason, there are various side effects, as cyberbullying has emerged as a serious issue afflicting children, adolescents and young adults. Machine learning techniques have conceivable ability to make automatic detection of bullying messages in social media, and this could develop a healthy and comparatively safe social media environment.

Keywords: Social Media; Big Data; NLP; Twitter; Machine Learning; Cyberbullying Detection.

I. INTRODUCTION

Opinion extraction, sentiment mining, affect analysis, analysis of emotion, review analyzing or mining, etc. are numerous comparative names with somewhat unique errands and with slightly different tasks. However, they are presently considered within the shades of opinion mining or sentiment analysis. This can be considered as an important application of NLP (Natural Language Processing).

Some of social media sites, includes social networking sites such as Facebook, MySpace, and Twitter; video sites such as YouTube; photo sharing such as Flicker, Photobucket, or Picasa; gaming sites and virtual worlds such as Kaneva, Club Penguin, Second Life, and the Sims; live casting such as Ustream or Twitch; instant messaging like Google talk, yahoo messenger or skype and blogs.

Authors in [35] and [12] use reciprocally the terms Big Social Data and Social Big Data to refer about the data which is generated by social media. To show the tremendous measure of data that is generated by social networking, [26] affirms that via Facebook (the most popular social media [11, 36, 31]), 10 million photos are getting uploaded almost every day. [21], focus attention to highlight the aspect that more than 250 million tweets are sent by Twitter every day consistently as well, also 3000+ photos are getting uploaded across Flicker every minute consistently without overlooking above 150 million web blogs posted daily. The expansion of Social Big Data is extremely valuable in numerous fields such as sociology, human science, psychology, governmental issues, politics and the very important commercial area. [16, 18]

Nonetheless, social media over and above have some consequential side effects as cyberbullying, that might have intense unfavorable impact and can transform the life of people, especially the children, adolescent and teenagers. Cyberbullying detection can be administrated from social media as it can be particularized as a supervised learning problem. A classfier is initially trained on a cyberbullying corpus that is labeled with mark by humans, and then later the learned classifier is then used as a result to recognize & perceive bullying message.[38] Three categories of information are commonly used inclusive of user demography, text, and social network features of cyberbullying detection [28]. Since the text content substance is the most definitely dependable as well as reliable, our work here spotlights on the text-based content cyberbullying detection.

II. SENTIMENT ANALYSIS APPROACHES

Sentiment Analysis or Opinion Mining intends to adjudge the viewpoint stance of a person with respect to any subject. There are two fundamental approaches used for opinion mining, those are:

- Machine Learning
- Semantic Orientation.

The machine learning approach is said to be belonging to supervised classification approach. This approach is more precise as each of the classifiers is initially trained on a collection of representative data which is known as corpus. In this way, it is called “supervised learning”. In a machine (supervised) learning based classification, two sorts of documents are required: training set as well as test set. A training set is utilized to prepare and learn the classifier and a test set is utilized to test the performance on execution of the automatic classifier. A huge number of machine learning techniques are available to access which classifies the opinions. Machine learning techniques like Naïve Bayes, Maximum Entropy (ME) and Support Vector Machines (SVM) have achieved extremely significant success in text categorization.[37]

III. PROCESS

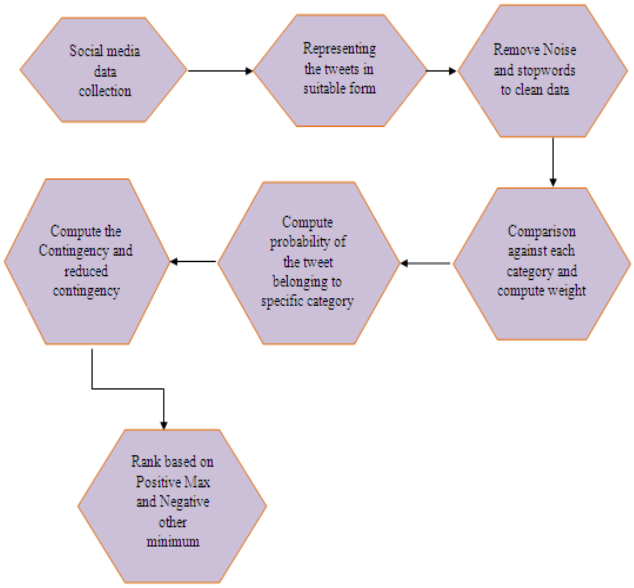

Fig I. Process to detect bullying messages

Fig I. shows the procedure followed:

1. Tweets for various cases acts as input
2. The tweets are collected from the twitter using OAuth API.
3. For each of the tweets the stop words are removed and clean tweets are also obtained.
4. Each tweet is then compared against the words belonging to 5 different categories.
5. Compute the probability using Naïve Bayes formula.
6. Compute the contingency
7. Compute the Reduced contingency
8. Rank the Tweets based on maximum positive and negative minimum

IV. LITERATURE REVIEW

A. Survey on Sentiment Analysis:

Paper [40], Nasukawa et al. demonstrates an approach to sentiment analysis in which they extract sentiments for particular subjects associated with its polarities i.e. negative or positive from a document, in spite of classifying the whole document.

Initial work on sentiment analysis focused on identifying polarity of reviews of product Opinions [23] and movie reviews from IMDB (Internet Movie data base) at entire document level [7].

Later work handles sentiment analysis at sentence level [27]. Recent studies are focus was shifted from sentence level to phrase-level [2] and short-text forms in response to the popularity of micro blogging services such as Twitter [6,1,17,3,5].

In [18], Ding et al. proposed an effective method for identifying semantic orientations of opinions expressed by reviewers on product features.

In sentiment analysis, Pandey et al. [37] presents a system, which can extract micro messages relevant to any specific topic from a blogging service such as Twitter and then analyze the messages to determine sentiments they carry and to classify them as neutral, positive or negative.

Table I. Review of Sentiment Analysis

References	Dataset	Features	Techniques	Classifications Approach
R Arora et al. [30]	Restaurant Reviews	Unigram, Bigrams, Trigrams	SVM, Naïve Bayes	Supervised
Y. Qe et al. [29]	Reviews to travel destination	Unigram Frequency	SVM, Naïve Bayes, character based N-gram model	Supervised
R. Prabowo et al. [34]	Movie reviews, Product reviews, MySpace Comments	POS tag, N-gram	SVM , Rule based Classifier	Supervised Learning and Rule-based Classification
E. Riloff et al. [14]	Movie Reviews, MPQA dataset	Unigram, bigram and extraction pattern feature	SVM	Supervised
C. Whitelaw et al. [8]	Movie Review	Adjective word frequency, percentage of appraisal groups	SVM	Supervised
P. D. Turney et al. [27]	Automobile bank, movie, travel reviews	adjectives and adverbs	PMI-IR	Unsupervised
A. Harb et al. [4]	Movie Reviews	adjectives and adverbs	Association Rule	Unsupervised

M. Taboada et al. [25]	Movie Reviews, Camera Reviews	Adjectives, Nouns, verbs, Adverbs, Intensifier, Negation	Dictionary based approach	Unsupervised
Ming Jiang et al. [24]	Chinese online reviews	Chinese words	LDA	unsupervised topic
William Claster et al. [41]	Tourism, travel and tweets	one-dimensional measure and Kohonen SOM for multiple characteristics	Naïve Bayes classifier	Supervised

B. Survey on social media big data:

According to indication in [35,12], the Social Big Data, produced represent all the data as well as information generated through the social media. This data is then perceived by: the substantial volume, the commotion that can be propagated (spam) and the dynamic characteristic property (the frequent changes day by day) [16]. They can likewise be perceived by an arrangement set of connections or links (due to connections between users), a structure or frame which is not aligned pattern in nature (as a result of the length of messages which is required by particularly some kind of micro-blogging, the existence of spelling inaccuracies or other) and the absence of fulfillment (as a result of satisfying the user requirements for privacy of their data) [15]. The mentioned attributes of Social Big Data qualify them unconventionally discrete from other version data on which elementary data mining techniques are enforcedly applied [16]. For mining such sort of data, various research issues, methodologies, approaches, procedures and techniques are derived.[18]

C. Survey on Effect of Social Media

Sentiment Analysis is used for the prediction of polarity of textual data into positive, negative and neutral classes. This textual data can be gathered from social media (e.g. twitter). Following Table 1[9] shows how social media creates its impact in various situations.

Table II. Social Media Effects

Reference	Paper	Description
Pontifícia Universidade Católica do Rio Grande do Sul	Can visualization techniques help journalists to deepen analysis of Twitter data?	In this paper the novel techniques shows how journalists and scholars can gain an improvised understanding about their audience before, during, and after an ebb and flow occasion.
El Noshokaty	Examine moral, emotional, and cognitive elements in petition language and determine their role in making online petitions successful	This paper shows that it is difficult to understand the associations between language (dialect), emotions (feelings), and behavior (conduct).
Dr. Shing Doong	Predicting Twitter hashtags popularity level	Author of this paper augments existing techniques that use information about the tweet author and time series arrangement analysis to likewise use wavelet transforms. These transforms in several cases enhance prediction precision by over 10%.
Zangerle, Schmidhammer and Specht	Analysing the usage of Wikipedia on Twitter: Understanding inter-language links	The authors find that the nature as well as quality of the articles tweeted is perpetually higher comparative to its alternatives, proposing that Twitter users regularly allude to an article which is of higher quality, regardless of the possibility that it is composed in another language.
Zimbra, Ghassi and Lee	Brand-related Twitter sentiment analysis using feature engineering and the dynamic architecture for artificial neural networks	Authors in this paper present an innovative technique for presenting the challenges related with the Twitter language and the review of mellow sentiment expressions. In addition both of these are concerned to the brand management practitioners.
Zeng, Starbird and Spiro	Rumors at the speed of light? Modeling the rate of rumor transmission during crisis	This paper estimate to measure the transmission speed of gossip avowing or rumor-affirming messages along with rumor-rectifying messages amid a prominent hostage crisis. The work hasessential ramifications for the developing field of crisis informatics.
team from Universidade Federal do ABC	The people have spoken: Conflicting Brazilian protests on Twitter	The analysis demonstrates in this paper how protesters efficiently arranged themselves systematically into groups, how the way they acted, also the utility of subject matter mining algorithms to extricate the notable conclusions, opinions, requests and demands of various groups.
Zhou, Zhao, Zhang and Wang	Measuring emotion bifurcation points for individuals in social media	This paper illustrates to what degree the emotions are diverged and suggests a novel framework to distinguish and visualize the nature of individuals' emotional divergence, then focuses on interpersonal organizations.

D. Survey on Cyberbullying

To originate a productive potent efficient practical cyberbullying identification model, we suggest expanding on discoveries within the psychology community. There are various reviews investigating the psychological dimensions of social collaborations that can be used to determine and analyse the cyberbullying risk elements, risk factors, risk aspects, which a model ought to consider. A large portion of the work in determining bullying among youth has concentrated on traditional bullying, conventional tormenting, or cyberbullying by the means of versatile mobile or visit chat-based scenes, e.g., [33, 10, 20, 32, 39, 19]. Previous contributions have concentrated different aspects of bullying and cyberbullying in their studies, e.g., whether guardians' viewpoint of adolescents' online conduct is causational with teenagers' vulnerability to cyberbullying [33,10], probabilities of exploitation and victimization [20] and passionate effect as well as emotional impact [32,39] in light of age and gender, and measuring the correlation between seriousness of online hostility and the quantity of spooks or bullies included [32]. While the outcomes about pervasiveness and determinants of cyberbullying change in the psychology literature, there are some critical patterns and regions of understanding [13,22] among these outcomes. [42]

V. CONCLUSION

Text mining is likewise alluded as textual data mining, which can be generally considered as text analytics. It refers to the process of deriving insights from text, include task such as sentiment analysis which suffers from problem of high dimensionality. This paper proposes the solution for text-based cyberbullying detection problem, where the vigorous, robust, potent and discriminative portrayals and distinctive appropriable representations of messages are critically analytic for an adequate detection system. By designing a semantic dropout noise, using stopwords, the development of auto-encoder is done as a specifically specialized representation of learning model for cyberbullying detection. The accomplishment of methodologies can be verified experimentally with the help of cyberbullying corpora from social medias like Twitter and MySpace. As a next stride the plan might further enhance the strength and robustness of the execution of learned advent by considering word arrangement order in the messages.

VI. ACKNOWLEDGEMENT

I wish to express my deepest sense of gratitude for immeasurably valuable guidance and support I have received from my supervisors Dr. Piyush Kumar Shukla, Mr. Uday Chourasia. I am very thankful to Dr. Rakesh Singhai, Coordinator, DDIPG, RGPV, Bhopal for providing support and facilities needed to accomplish the work. I express my deep gratitude to God and last but not least my family for love, inspiration and moral support.

VII. REFERENCES

[1] A. Agarwal, B. Xie, I. Vovsha, O. Rambow and R. Passonneau, "Sentiment Analysis of Twitter Data," In Proceedings of workshop on languages of social media, 2011. pp. 30-38.

[2] A. Agarwal, F. Biadsy, and K. R. Mckeown. "Contextual phrase-level polarity analysis using lexical affect scoring and syntactic n-grams." Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics. Association for Computational Linguistics, 2009. pp. 24-32.

[3] A. Go, R. Bhayani, L. Huang, "Twitter sentiment classification using distant supervision", CS224N Project Report, Stanford, 2009. pp. 1-12.

[4] A. Harb, M. Plantié, G. Dray, M. Roche, F. Trousset, and P. Poncelet. "Web Opinion Mining: How to extract opinions from blogs?" In Proceedings of the 5th international conference on Soft computing as transdisciplinary science and technology, ACM, 2008. pp. 211-217.

[5] A. Pak, and P. Paroubek "Twitter as a corpus for sentiment analysis and opinion mining," Proceedings of the seventh International Conference on Language Resource and Evolution (LREC'10), 2010. pp. 19-21.

[6] B. O'Connor, R. Balasubramanyan, B. R. Routledge, and N. A. Smith. "From tweets to polls: Linking text sentiment to public opinion time series." ICWSM 11, no. 122-129 ,2010. pp. 1-2.

[7] B. Pang, L. Lee, and S. Vaithyanathan. "Thumbs up? : sentiment classification using machine learning techniques." Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10. Association for Computational Linguistics, 2002. pp. 79-86.

[8] C. Whitelaw, N. Garg, and S. Argamon "Using appraisal groups for sentiment analysis," presented at the Proceedings of the 14thACM international conference on Information and knowledge management, Bremen, Germany, 2005.

[9] D. M. Haughton, J. J. Xu, D. J. Yates, X. Yan "Introduction to Data Analytics and Data Mining for Social Media Minitrack", 49th Hawaii International Conference on System Sciences, 2016.p. 1414.

[10] D. M. Law, J. D. Shapka, and B. F. Olson. "To control or not to control? Parenting behaviours and adolescent online aggression." Computers in Human Behavior 26.6,2010 pp.1651-1656.

[11] D. Quinn, C. Liming, and M. Maurice "Does age make a difference in the behaviour of online social network users?" Internet of Things (iThings/CPSCom), 2011 International Conference on and 4th International Conference on Cyber, Physical and Social Computing. IEEE, 2011.

[12] D. T. Nguyen, D. Hwang, and J. J. Jung "Time-frequency social data analytics forunderstanding social bigdata", Intelligent Distributed Computing VIII, pages 223–228. Springer, 2015.

[13] D. Wolke, T. Lereya, and N. Tippett. "Individual and Social Determinants of Bullying and Cyberbullying." Cyberbullying. Ed. T. Vollink, Ed. F. Dehue, Ed. C. Guckin. Routledge, 2016. 26-53.

[14] E. Riloff, S. Patwardhan, and J.Wiebe "Feature subsumption for opinion analysis." presented at the Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing, Sydney, Australia, 2006.

[15] F. Colace, M. De Santo, and L. Greco. "A probabilistic approach to tweets' sentiment classification." Affective computing and intelligent interaction (ACII), 2013 humaine association conference on. IEEE, 2013.

[16] G. Barbier, and H. Liu. "Data mining in social media." Social network data analytics ,2011: pp. 327-352.

[17] H. Saif, Y. He, and H. Alani “Semantic Sentiment Analysis of Twitter”, Proceedings of 11th International Conference on Semantic Web.

[18] I. Guellil and K. Boukhalfa. "Social big data mining: a survey focused on opinion mining and sentiments analysis." Programming and Systems (ISPS), 2015 12th International Symposium on. IEEE, 2015. pp. 1-10.

[19] J. J. Dooley, J. Pyżalski, and D. Cross. "Cyberbullying versus face-to-face bullying: A theoretical and conceptual review." Zeitschrift für Psychologie/Journal of Psychology 217.4 ,2009. pp.182-188.

[20] J. Piazza and J. M. Bering. "Evolutionary cyber-psychology: Applying an evolutionary framework to Internet behavior." Computers in Human Behavior 25.6 ,20009. pp.1258-1269.

[21] J. Tang, Y. Chang, and H. Liu “Mining social media with social theories: A survey”, ACM SIGKDD Explorations Newsletter 15.2, June 2014. pp.20-29.

[22] J. W. Patchin and S. Hinduja. “Cyberbullying - An Update and Synthesys of the Research.” Cyberbullying Prevention and Response.: Expert perspectives, 2012. 13-35.

[23] M. Hu and B. Liu. "Mining and summarizing customer reviews." Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2004. pp. 168-177.

[24] M. Jiang, J. Wang, X. Wang, J. Tang and C. Wu. ‘A mixture language model for the classification of Chinese online reviews’, Int. J. Information and Communication Technology, Vol. 7, No. 1, pp.109–122.

[25] M. Taboada, J. Brooke, M. Tofiloski, K. Voll, and M. Stede, "Lexicon-based methods for sentiment analysis," Computer. Linguist. vol. 37, pp. 267-307, 2011.

[26] N. Jaafar, M. Al-Jadaan, and R. Alnutaifi. "Framework for social media big data quality analysis." In New Trends in Database and Information Systems II, Springer International Publishing, 2015. pp. 301-314.

[27] P. D. Turney "Thumbs up or thumbs down? : semantic orientation applied to unsupervised classification of reviews." Proceedings of the 40th annual meeting on association for computational linguistics. Association for Computational Linguistics, July 2002. pp. 417-424.

[28] Q. Huang, V. K. Singh, and P. K. Atrey, “Cyber bullying detection using social and textual analysis,” Proceedings of the 3rd International Workshop on Socially-Aware Multimedia. ACM, 2014, pp. 3–6.

[29] Q. Ye, Z. Zhang, and R. Law "Sentiment classification of online reviews to travel destinations by supervised machine learning approaches," Expert Systems with Applications, vol. 36, 2009. pp. 6527-6535.

[30] R. Arora and S. Srinivasa “Communication Systems and Networks (COMSNETS), 2014 Sixth International Conference on. IEEE, 2014.pp.1-6.

[31] R. Farahbakhsh, X. Han, A. Cuevas, and N. Crespi. "Analysis of publicly disclosed information in facebook profiles." Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining. ACM, 2013.

[32] R. Ortega, P. Elipe, J. A. Mora-Merchán, J. Calmaestra, and E. Vega "The emotional impact on victims of traditional bullying and cyberbullying: A study of Spanish adolescents." Zeitschrift für Psychologie/Journal of Psychology 217.4 ,2009. pp.197-204.

[33] R. P. Ang, W. H. Chong, S. Chye, & V. S. Huan "Loneliness and generalized problematic Internet use: Parents’ perceived knowledge of adolescents’ online activities as a moderator." Computers in Human Behavior 28.4 ,2012. pp. 1342-1347.

[34] R. Prabowo and M. Thelwall "Sentiment analysis: A combined approach," Journal of Informetrics, vol. 3, 2009. pp. 143-157.

[35] R. R. Mukkamala, A..Hussain, and R.Vatrapu “Fuzzy-set based sentiment analysis of big social data”, Enterprise Distributed Object Computing Conference (EDOC), 2014 IEEE 18th International, pages 71–80, Sept 2014.

[36] R. Rabade, N. Mishra, and S. Sharma. "Survey of influential user identification techniques in online social networks." Recent advances in intelligent informatics. Springer International Publishing, 2014. pp. 359-370.

[37] R. Sharma, S. Nigam, R. Jain, “Supervised Opinion Mining Techniques: A Survey”, International Journal of Information and Computation Technology, 2013.

[38] R. Zhao and K. Mao. "Cyberbullying Detection based on Semantic-Enhanced Marginalized Denoising Auto-Encoder." IEEE Transactions on Affective Computing ,2016.

[39] T. E. Waasdorp, and C. P. Bradshaw. "Examining student responses to frequent bullying: a latent class approach." Journal of Educational Psychology 103.2,2011. p.336.

[40] T. Nasukawa and J. Yi. "Sentiment analysis: Capturing favorability using natural language processing." Proceedings of the 2nd international conference on Knowledge capture. ACM, 2003. pp.70–77.

[41] W. Claster, P. Pardo, M. Cooper. and K. Tajeddini, ‘Tourism, travel and tweets: algorithmic text analysis methodologies in tourism’, Middle East J. Management, Vol. 1, No. 1,2013. pp.81–99.

[42] Y. N. Silva, C. Rich, D. Hall “BullyBlocker: Towards the Identification of Cyberbullying in Social Networking Sites”, International Conference on Advances in Social Networks Analysis and Mining (ASONAM), 2016 IEEE/ACM.