International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

PREDICTING CYBER BULLYING ON SOCIAL MEDIA IN THE BIG DATA ERA USING EXTREME LEARNING MACHINE

Dr. V Chandra Sekhar[1], Chintalapati Sindhu Sri[2]

#1: Associate professor and Head of the Department of Computer Science and Technology

#2: M.Tech Scholar Computer science and Technology

[#1 #2]Sagi Rama Krishnam Raju Engineering College Bhimavaram

Abstract

Cyber bullying disturbs harassment online, with alarming implications. It exists in different ways, and is in textual format in most social networks. There is no question that over 1.96 billion of them would have an inescapable social operation. However, the developing decade presents genuine difficulties and the online-conduct of clients have been put to address. Expanding instances of provocation and harassing alongside instances of casualty has been a difficult issue. Programmed discovery of such episodes requires smart frameworks. A large portion of the current studies have been moving towards this issue with standard machine learning models and most of the models produced in these studies are scalable at one time into a solitary social network. Deep learning-based models have discovered ways in the identification of digital harassing occurrences, asserting that they can beat the restrictions of the ordinary models, and improve the discovery execution using extreme machine learning techniques. However, numerous old-school models are accessible to control the incident, the need to successfully order the tormenting is as yet weak. To successfully screen the harassing in the virtual space and to stop the savage outcome with the execution of Machine learning and Language preparing. A system is proposed to give a double characterization of cyberbullying. Our technique utilizes an inventive idea of CNN for content examination anyway the current strategies utilize a guileless way to deal with furnish the arrangement with less precision. A current dataset is utilized for experimentation and our system is proposed with other existing methods and is found to give better precision and grouping.

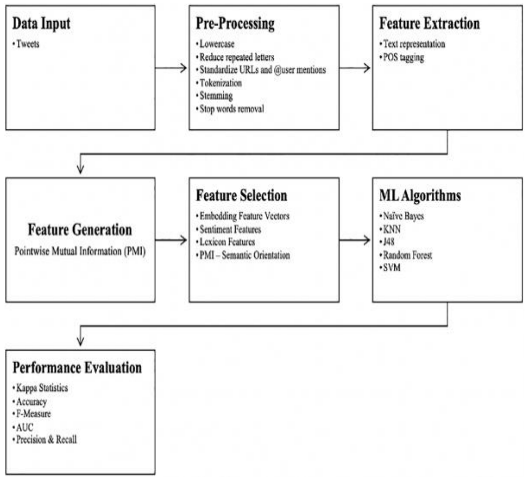
Introduction

In this article, we propose a cyberbullying detection framework to generate features from Twitter content(tweets) by leveraging a pointwise mutual information technique. Based on these features, we have developed a supervised machine learning solution for cyberbullying detection and multi-class categorization of its severity in

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

Twitter. We have applied Embedding, Sentiment, and Lexicon features along with PMI-semantic orientation. Extracted features were applied
with Naïve Bayes, KNN, Decision Tree, Random Forest, and Support Vector Machine algorithms. In this article we first briefly present background on key areas that our study focuses upon. In section 2, we outline related work in the state of the art related to classification of severity of cyberbullying. Section 3 provides the background for data usage for cyberbullying detection and its accessibility. Section 4 and 5 provide the research methodology framework used for cyberbullying detection and its severity. Proposed framework evaluation and results are presented in section 6 and comparison of baseline and proposed framework results are provided in section 7. Finally, the article provides some conclusions related to the significance of the proposed framework and suggests some future work.

System Architecture:



Literature Review:

Cyberbullying takes various forms, such as circulating filthy rumours on the bases of racism, gender, disability, religion and sexuality; humiliating a person; social exclusion; stalking; threatening someone online; and displaying personal information about an individual that was shared in confidence [1].According to the national advocacy group in US, the bullying can take several forms: racism and sexuality are two of these [2]. Based on a report at Pew Research Centre, two distinct categories of online harassment have been described among internet users. The first category includes less severe experiences: it involves swearing and humiliation, because those who see or experience it often claim they ignore it. The second category of harassment although targeting a smaller number of online users, includes more severe experiences such as physical threats, longterm harassment, trapping and sexual harassment [3].Assessing the severity level of a cyberbullying incident may be important in depicting the different correlations observed in cyberbullying victims, and principally, how these incidents impact victims' experience with cyberbullying [4]. Researchers, however, have not paid enough attention to the extent to which the different cyberbullying incidents could have more severe impact upon victims. Therefore, it is significant to develop a method to identify the severity of cyberbullying in OSNs.

Problem Definition:

This paperhighlights the limitation of existing techniques related to cyberbullying detection and its severity levels.

We provide a systemic framework for identifying cyberbullying severity in online social networks, which is based on previous research from different disciplines. We build machine learning multi-classifier for classifying cyberbullying severity into different levels. Our cyberbullying detection model work with multi-class classification problem and as well as for binary class classification problem

Pre-processing step

The collected data was pre-processed before assigning severity levels. Tweets were converted to lower case to avoid any sparsity issue, reduced repeated letters, standardized URLs and @usermention to remove noise in the tweets. Tokenization was applied with Twitter-specific tokenizer based on the CMU TweetNLP library [8] and only words with minimum frequency of 10 were kept. Tokenization is the process of breaking a text corpus up into most commonly words, phrases, or other meaningful elements, which are then called tokens. Finally, stop-words and stemming procedures were performed before feature extraction. Stop words are defined as the insignificant words that appear in document which are not specific or discriminatory to the different classes. Stemming refers to the process of reducing words to their stems or roots. For instance, singular, plural and different tenses are consolidated into a single word. We applied stemming with an iterated version of the Lovins stemmer, it stems the word until it no further changes prior to extracting topic model features [9].

Feature extraction step

All tweets were represented with bag-of-words which is one of the most appropriate and quickest approaches. In this approach, text is represented by set of words and

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

each word is treated as an independent feature. We applied part-of-speech (POS) tagging with Twitter-specific tagger based on the CMU TweetNLP library [6] for word sense disambiguation. The POS tagger assigns part-of-speech tag to each word of the given text in the form of tuples (*word*, *tag*), for instance, noun, verb, adjectives, etc.

Feature generation step

We applied document level classification and measured semantic orientation of each word in the corpus. In the document level classification, phrases were extracted using the POS tags. Once phrases have been extracted from the dataset, then their semantic orientation in terms of either cyberbullying or non-cyberbullying was determined. In order to achieve this goal, the concept of pointwise mutual information (PMI) [5] was used to calculate the semantic orientation for each word in a corpus of tweets. The PMI between two words, *word1* and *word2*, is defined as follows:

$$PMI(word_1, word_2) = \log_2\left[\frac{p(word_1 \& word_2)}{p(word_1)p(word_2)}\right]$$

The score was calculated by subtracting the PMI of the target word with a cyberbullying class from the PMI of the target word with a non-cyberbullying class. This method was clearly well suited for domain specific lexicon generation with PMI score, so we created our domain specific lexicon with PMI semantic orientation for each word and phrase by using Turney's technique [6]. Semantic Orientation of phrase, *phrase* is calculated as follows:

$$SO(phrase) = PMI(phrase, \text{"non} - \text{cyberbullying"}) - PMI(phrase, \text{"cyberbullying"})$$

Turney's method provides a representative lexicon-based technique consisting of three steps. First, phrases are extracted from the dataset. Second, sentiment polarity is estimated using PMI of each extracted phrase, which measures the statistical dependency between two terms. Lastly, polarity of all phrases in dataset is averaged out as its sentiment polarity. Turney's PMI technique does not depend on hard-coded semantic rules, so users may readily apply the technique into different contexts [5].

Feature engineering and selection step

Feature engineering is the process of generating or deriving features from raw data or corpus. Creation of additional features inferring from existing features is known as feature engineering [6]. It is not the number of features, but the quality of features that are fed into machine learning algorithm that directly affects the outcome of the model prediction [7].

One of the most common approaches to improve cyberbullying detection is to perform feature engineering, and most common features that improve quality of cyberbullying detection classifier performance are; textual, social, user, sentiment, word embeddings features [8]. Since social and user features were not available in the dataset provided by [12], we attempted to build features based on the textual context and their semantic orientation. As a consequence, we propose the following features to improve cyberbullying detection in multi-class classification setting for detecting cyberbullying predefined severity as well as same approach for the binary classification setting (whether or not cyberbullying behaviour exists in the tweets).

The following feature types were applied after pre-processing:

Embedding Feature Vector: In this study, tweet-level feature representation using pre-trained Word2Vec embeddings were applied. We used 400-dimension embeddings of 10 million tweets from the Edinburgh corpus [9].

Sentiment Feature Vector: SentiStrength [10] was used to calculate positive and negative score of each tweet.

Lexicon Feature Vector: Multiple phrase level lexicons were applied in this study that identify positive and negative contextual polarity of sentiment expression in our dataset. Lexicons includes: MPQA Subjectivity Lexicon [11], BingLiu [12], AFINN, Sentiment-140 , Expanded NRC-10, NRC Hashtag Sentiment lexicon , SentiWordnet , NRC-10 , and NRC Hashtag Emotion Association Lexwicon [19].

PMI-Semantic Orientation: In doing so, we processed previously generated domain specific lexicon (section 4.4) which contained mutual information of each word in the corpus. This PMI input approach assigns a PMI score to each word in the document. PMI-Semantic Orientation is then calculated for each document by subtracting the PMI of the target word.

Dealing with class imbalance data

Class imbalance refers to the scenario where the number of instances from one class is significantly greater than that of another class . Most machine learning algorithms work best when the number of instances of each of the classes are roughly equal. However, in many real-life applications and non-synthetic datasets, the data is imbalanced; that is, an important class (usually referred to as the minority class) may have many fewer samples than the other class (usually referred to as the majority class). In such cases, standard classifiers tend to be overwhelmed by the large class and ignore the small distributed instances. It usually produces a biased classifier that has higher predictive accuracy over majority classes, but poorer predictive accuracy over minority class. One way of solving the imbalanced class problem is to modify the class distributions in the training data by over-sampling the minority class or under sampling the majority class. SMOTE (Synthetic Minority Over-sampling Technique) [22] is specifically designed for learning from imbalanced datasets and is one of the most adopted approaches to deal with class imbalance due to its simplicity and effectiveness. It is a combination of oversampling and under sampling.

Our data set turned out to have an imbalanced class distribution (as shown in Table 3), that is, cyberbullying tweets with high severity class distribution were 4%, Medium 6%, Low 3%, and non-cyberbullying class distribution having 87%. Accordingly, we employed the SMOTE over sampling technique for our study. The next section presents the comparative results before and after using each machine learning approach.

Machine learning algorithms selection step

Choosing the best classifier is the most significant phase of the text classification pipeline. We cannot efficiently determine the most effective model for a text classification implementation without a full conceptual comprehension of each algorithm. The features (given in 4.E section) obtained from the tweets have been used to build a model to detect cyberbullying behaviours and its severity. In order to select the best classifier, we tested several machine learning algorithms namely: Naïve Bayes, Support Vector Machine (SVM), Decision Tree, Random Forest, and K-Nearest Neighbours (KNN).

Naïve bayes

In the field of machine learning, Naïve Bayes [21] is regarded as one of the most efficient and effective inductive learning algorithms and has been used as an effective classifier in several social media studies [38]. Since 1950s, Naïve Bayes classification for text has been commonly used in document categorization assignments and has ability to classify any type of data from text, network features, phrases, and so on. This technique is a generative model, it refers to how dataset

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

is generated based on probabilistic model. By sampling from this model, it can generate new data similar to the data on which the model is being trained [22]. In our study, we used the most basic version of Naïve Bayes classifier for textual features and word embeddings.

K-Nearest Neighbours (KNN)

The K-Nearest Neighbours (KNN) is a supervised learning algorithm and one of the simplest instance-based learning algorithms suitable for multi-class problems [23]. In this algorithm, distance is used to classify a new sample from its neighbour. Thus, finds the K-nearest neighbours among the training set and places an object into the class that is most frequent among its k nearest neighbours. KNN is considered as non-parametric lazy learning algorithm that does not make any assumptions on the underlying data distribution.

Decision trees (J48)

In machine learning, decision tree is one of the well-known classification algorithms and one of the most widely used inductive learning method. It can handle training data with missing values and can handle both continuous and discrete attributes. Decision trees are built from labelled training data using the concept of information entropy [16]. Their robustness to noisy data and their capability to learn disjunctive expressions seem suitable for text classification [18].

Random forest

Random forest (RF) is an ensemble algorithm which is used for the classification and regression problem. RF creates several decision trees classifiers on a random subset of data samples and features. The classification of new sample is done by majority voting of decision trees. The main advantage of RF is that it runs efficiently on large datasets, it is an effective method for estimating missing data, and offers good accuracy even if a large portion of the data is missing [19].

Support Vector Machine (SVM)

SVM is a pattern recognition supervised learning algorithm to classify both linear and non-linear data. The primary concept of SVM is to determine separators that can best distinguish the distinct classes in the search space. The data points that separate one or more hyperplane using essential training tuples are called support vectors.

In a few cases, nonlinear SVM classifier is used when all the data points cannot be separated by a straight line. Nonlinear function generally uses the kernel function namely; linear kernels, polynomial kernel, RBF kernel, and sigmoid kernel are the popular kernels. Normally, Radial basis function (RBF) kernel performs better than others when the number of features is much lower than the number of observations and Polynomial kernels works better when the data is normalized [20]. In order to achieve high classification performance, it is necessary to properly select kernel parameters. In this study, we selected RBF and Polynomial kernel. SVM is traditionally used for binary classification and it needs to be modified to work with multi-class classification since we have considered four classes for cyberbullying severity detection. There are two different types of techniques to tackle this problem; i) One-against-one: In this technique, SVM combines several binary classifiers, ii) one-against-all: In this technique, SVM considers all data at once [29].

By training our SVM model, each of the four classes high, medium, low and non-cyberbullying were applied as target variables using one-against-all approach. This strategy consists of fitting one classifier per class. For each classifier, the class is fitted against all the other classes

Conclusion

The use of internet and social media has clear advantages for societies, but their frequent use may also have significant adverse consequences. This involves unwanted sexual exposure, cybercrime and cyberbullying. We developed a model for detecting cyberbullying behaviour and its severity in Twitter. Feature generation with PMI at pre-processing stage has proven to be the efficient technique to handle class imbalance in binary and multi-class classification where misclassification for minority class (es) has higher cost in terms of its impact on reliability of detection model. The developed model is a feature-based model that uses features from tweets contents to develop a machine learning classifier for classifying the tweets as cyberbullying or non-cyberbullying and its severity as low, medium, high or none

References

1.Fire M, Goldschmidt R, Elovici Y. Online Social Networks: Threats and Solutions. IEEE Commun Surv Tutor. 2014;16: 2019–2036.

2.Penni J. The future of online social networks (OSN): A measurement analysis using social media tools and application. Telemat Inform. 2017;34: 498–517.

3.Lauw H, Shafer JC, Agrawal R, Ntoulas A. Homophily in the Digital World: A LiveJournal Case Study. IEEE Internet Comput. 2010;14: 15–23.

4.Rezvan M, Shekarpour S, Balasuriya L, Thirunarayan K, Shalin VL, Sheth A. A Quality Type-aware Annotated Corpus and Lexicon for Harassment Research. Proceedings of the 10th ACM Conference on Web Science. New York, NY, USA: ACM; 2018. pp. 33–36.

5.Hee CV, Jacobs G, Emmery C, Desmet B, Lefever E, Verhoeven B, et al. Automatic detection of cyberbullying in social media text. PLOS ONE. 2018;13: e0203794. pmid:30296299

6.Hosseinmardi H, Shaosong Li, Zhili Yang, Qin Lv, Rafiq RI, Han R, et al. A Comparison of Common Users across Instagram and Ask.fm to Better Understand Cyberbullying. 2014 IEEE Fourth International Conference on Big Data and Cloud Computing. 2014. pp. 355–362.

7.Citron DK. Addressing Cyber Harassment: An Overview of Hate Crimes in Cyberspace. the Internet. 2015;6: 12.

8.Wall D. What are Cybercrimes? Crim Justice Matters. 2004;58: 20–21.

9.Abu-Nimeh S, Chen T, Alzubi O. Malicious and Spam Posts in Online Social Networks. Computer. 2011;44: 23–28.

10.Doerr B, Fouz M, Friedrich T. Why Rumors Spread So Quickly in Social Networks. Commun ACM. 2012;55: 70–75.

11.Ferrara P, Ianniello F, Villani A, Corsello G. Cyberbullying a modern form of bullying: let's talk about this health and social problem. Ital J Pediatr. 2018;44. pmid:29343285

12.Volk AA, Veenstra R, Espelage DL. So you want to study bullying? Recommendations to enhance the validity, transparency, and compatibility of bullying research. Aggress Violent Behav. 2017;36: 34–43.

13.Sampasa-Kanyinga H, Roumeliotis P, Xu H. Associations between Cyberbullying and School Bullying Victimization and Suicidal Ideation, Plans and Attempts among Canadian Schoolchildren. PLOS ONE. 2014;9: e102145. pmid:25076490

14.Safaria T. Prevalence and Impact of Cyberbullying in a Sample of Indonesian Junior High School Students. Turk Online J Educ Technol. 2016;15: 10.

15.Anderson T, Sturm B. Cyberbullying: From Playground to Computer. Young Adult Libr Serv. 2007;5: 24.

16.Bauman S, Toomey RB, Walker JL. Associations among bullying, cyberbullying, and suicide in high school students. J Adolesc. 2013;36: 341–350. pmid:23332116

17.Foody M, Samara M, Carlbring P. A review of cyberbullying and suggestions for online psychological therapy. Internet Interv. 2015;2: 235–242.

18.Fridh M, Lindström M, Rosvall M. Subjective health complaints in adolescent victims of cyber harassment: moderation through support from parents/friends—a Swedish population-based study. BMC Public Health. 2015;15: 949. pmid:26399422

19.Gini G, Espelage DL. Peer Victimization, Cyberbullying, and Suicide Risk in Children and Adolescents. JAMA. 2014;312: 545–546. pmid:25096695

20.Nixon CL. Current perspectives: the impact of cyberbullying on adolescent health. Adolesc Health Med Ther. 2014;5: 143–158. pmid:25177157

21.Myers C-A, Cowie H. Cyberbullying across the Lifespan of Education: Issues and Interventions from School to University. Int J Environ Res Public Health. 2019;16. pmid:30987398

22.Duggan M. Online Harassment 2017. In: Pew Research Center: Internet, Science & Tech [Internet]. 11 Jul 2017 [cited 18 Aug 2019]. https://www.pewinternet.org/2017/07/11/online-harassment-2017/.

23.Duggan M. Online Harassment. In: Pew Research Center: Internet, Science & Tech [Internet]. 22 Oct 2014 [cited 19 Aug 2019]. https://www.pewinternet.org/2014/10/22/online-harassment/.

24.Camacho S, Hassanein K, Head M. Understanding the Factors That Influence the Perceived Severity of Cyber-bullying. In: Nah FF-H, editor. HCI in Business. Springer International Publishing; 2014. pp. 133–144.

25.Reynolds K, Kontostathis A, Edwards L. Using Machine Learning to Detect Cyberbullying. 2011 10th International Conference on Machine Learning and Applications and Workshops. 2011. pp. 241–244.

26.Potha N, Maragoudakis M. Cyberbullying Detection using Time Series Modeling. 2014 IEEE International Conference on Data Mining Workshop. 2014. pp. 373–382.

27.Einarsen S, Hoel H, Cooper C. Bullying and Emotional Abuse in the Workplace: International Perspectives in Research and Practice. CRC Press; 2002.

28.Dadvar M, de Jong F. Cyberbullying detection: a step toward a safer internet yard. Proceedings of the 21st international conference companion on World Wide Web—WWW '12 Companion. Lyon, France: ACM Press; 2012. p. 121.

29.Zuckerberg M. One Billion People on Facebook. In: One Billion People on Facebook [Internet]. 2012 [cited 20 Oct 2019]. https://newsroom.fb.com/news/2012/10/one-billion-people-on-facebook/.

30.Kurka DB, Godoy A, Von Zuben FJ. Online Social Network Analysis: A Survey of Research Applications in Computer Science. ArXiv150405655 Phys. 2015 [cited 24 Aug 2019]. http://arxiv.org/abs/1504.05655.

31.Bayzick J, Kontostathis A, Edwards L. Detecting the Presence of Cyberbullying Using Computer Software. 2011.

32.Dinakar K, Reichart R, Lieberman H. Modeling the Detection of Textual Cyberbullying. 2011; 7.

33.Ashktorab Z. A Study of Cyberbullying Detection and Mitigation on Instagram. CSCW Companion. 2016.

34.Chavan VS, Shylaja S S. Machine learning approach for detection of cyber-aggressive comments by peers on social media network. 2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI). 2015. pp. 2354–2358.

35.Van Hee C, Lefever E, Verhoeven B, Mennes J, Desmet B, De Pauw G, et al. Detection and Fine-Grained Classification of Cyberbullying Events. Proceedings of the International Conference Recent Advances in Natural Language Processing. Hissar, Bulgaria: INCOMA Ltd. Shoumen, BULGARIA; 2015. pp. 672–680. https://www.aclweb.org/anthology/R15-1086.

36.Nalini K, Sheela LJ. Classification of Tweets Using Text Classifier to Detect Cyber Bullying. In: Satapathy SC, Govardhan A, Raju KS, Mandal JK, editors. Emerging ICT for Bridging the Future—Proceedings of the 49th Annual Convention of the Computer Society of India CSI Volume 2. Springer International Publishing; 2015. pp. 637–645.

37.Jaidka K, Ahmed S, Skoric M, Hilbert M. Predicting elections from social media: a three-country, three-method comparative study. Asian J Commun. 2019;29: 252–273.

38.Al-garadi MA, Varathan KD, Ravana SD. Cybercrime Detection in Online Communications. Comput Hum Behav. 2016;63: 433–443.
39.Kavanaugh AL, Fox EA, Sheetz SD, Yang S, Li LT, Shoemaker DJ, et al. Social media use by government: From the routine to the critical. Gov Inf Q. 2012;29: 480–491.
40.Xu J-M, Jun K-S, Zhu X, Bellmore A. Learning from Bullying Traces in Social Media. Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Montréal, Canada: Association for Computational Linguistics; 2012. pp. 656–666. https://www.aclweb.org/anthology/N12-1084.
41.Zhao R, Zhou A, Mao K. Automatic Detection of Cyberbullying on Social Networks Based on Bullying Features. Proceedings of the 17th International Conference on Distributed Computing and Networking. New York, NY, USA: ACM; 2016. p. 43:1–43:6.
42.Gimpel K, Schneider N, O'Connor B, Das D, Mills D, Eisenstein J, et al. Part-of-Speech Tagging for Twitter: Annotation, Features, and Experiments. Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies. Portland, Oregon, USA: Association for Computational Linguistics; 2011. pp. 42–47. https://www.aclweb.org/anthology/P11-2008.
43.Lovins JB. Development of a stemming algorithm. Mech Transl Comp Linguist. 1968;11: 22–31.
44.Turney P. Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews. Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics. Philadelphia, Pennsylvania, USA: Association for Computational Linguistics; 2002. pp. 417–424.
45.Garrett M, Kuiper P, Hood K, Turner D. Leveraging Mutual Information to Generate Domain Specific Lexicons. 2018; 7.
46.Pattnaik PK, Rautaray SS, Das H, Nayak J. Progress in Computing, Analytics and Networking: Proceedings of ICCAN 2017. Springer; 2018.
47.Mehta R. Big Data Analytics with Java. Packt Publishing Ltd; 2017.
48.Rosa H, Pereira N, Ribeiro R, Ferreira PC, Carvalho JP, Oliveira S, et al. Automatic cyberbullying detection: A

systematic review. Comput Hum Behav. 2019;93: 333–345.
49.Petrović S, Osborne M, Lavrenko V. The Edinburgh Twitter Corpus. Proceedings of the NAACL HLT 2010 Workshop on Computational Linguistics in a World of Social Media. Los Angeles, California, USA: Association for Computational Linguistics; 2010. pp. 25–26.
https://www.aclweb.org/anthology/W10-0513.
50.Thelwall M, Buckley K, Paltoglou G. Sentiment strength detection for the social web. JASIST. 2012;63: 163–173.
51.Wilson T, Wiebe J, Hoffmann P. Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis. 8.
52.Hu M, Liu B. Mining and Summarizing Customer Reviews. 2014; 10.
53.Nielsen FÅ. A new ANEW: Evaluation of a word list for sentiment analysis in microblogs. ArXiv11032903 Cs. 2011 [cited 17 Sep 2019]. http://arxiv.org/abs/1103.2903.
54.Mohammad S, Kiritchenko S, Zhu X. NRC-Canada: Building the State-of-the-Art in Sentiment Analysis of Tweets. Second Joint Conference on Lexical and Computational Semantics (*SEM), Volume 2: Proceedings of the Seventh International Workshop on Semantic Evaluation (SemEval 2013). Atlanta, Georgia, USA: Association for Computational Linguistics; 2013. pp. 321–327.
https://www.aclweb.org/anthology/S13-2053.
55.Bravo-Marquez F, Frank E, Mohammad SM, Pfahringer B. Determining Word-Emotion Associations from Tweets by Multi-label Classification. 2016 IEEE/WIC/ACM International Conference on Web Intelligence (WI). Omaha, NE, USA: IEEE; 2016. pp. 536–539.
56.Kiritchenko S, Zhu X, Mohammad SM. Sentiment Analysis of Short Informal Texts. J Artif Intell Res. 2014;50: 723–762.

57.Baccianella S, Esuli A, Sebastiani F. SENTIWORDNET 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining. 2010; 5.
58.Mohammad SM, Turney PD. Crowdsourcing a Word-Emotion Association Lexicon. ArXiv13086297 Cs. 2013 [cited 17 Sep 2019]. http://arxiv.org/abs/1308.6297.
59.Mohammad SM, Kiritchenko S. Using Hashtags to Capture Fine Emotion Categories from Tweets.Comput Intell. 2013; 22.

60.Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: Synthetic Minority Over-sampling Technique. J Artif Intell Res. 2002;16: 321–357.

61.Ng AY, Jordan MI. On Discriminative vs. Generative Classifiers: A comparison of logistic regression and naive Bayes. In: Dietterich TG, Becker S, Ghahramani Z, editors. Advances in Neural Information Processing Systems 14. MIT Press; 2002. pp. 841–848. http://papers.nips.cc/paper/2020-on-discriminative-vs-generative-classifiers-a-comparison-of-logistic-regression-and-naive-bayes.pdf.

62.Foster D. Generative Deep Learning: Teaching Machines to Paint, Write, Compose, and Play. O'Reilly Media, Inc.; 2019.

63.Witten Ian H., Frank Eibe, Hall Mark A. Data Mining: Practical Machine Learning Tools and Techniques. Elsevier; 2011.

64.Quinlan JR. C4.5: Programs for Machine Learning. Morgan Kaufmann; 1993.

65.Li YH, Jain AK. Classification of Text Documents. Comput J. 1998;41: 10.

66.Awad M, Khanna R. Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers. Apress; 2015.

67.Abraham A. Emerging Technologies in Data Mining and Information Security. Springer; 2018.

68.Yi Liu, Zheng YF. One-against-all multi-class SVM classification using reliability measures. Proceedings 2005 IEEE International Joint Conference on Neural Networks, 2005. Montreal, Que., Canada: IEEE; 2005. pp. 849–854.

69.Alber M, Zimmert J, Dogan U, Kloft M. Distributed optimization of multi-class SVMs. PLOS ONE. 2017;12: e0178161. pmid:28570703

70.Kowsari K, Meimandi KJ, Heidarysafa M, Mendu S, Barnes LE, Brown DE. Text Classification Algorithms: A Survey. Information. 2019;10: 150.

71.Sokolova M, Lapalme G. A systematic analysis of performance measures for classification tasks. Inf Process Manag. 2009;45: 427–437.

72.Lagopoulos A, Kapraras N, Amanatiadis V, Fachantidis A, Tsoumakas G. Classifying Biomedical Figures by Modality via Multi-Label Learning. IEEE J Biomed Health Inform. 2019; 1–1. pmid:30835232

73.Huang J, Ling CX. Using AUC and Accuracy in Evaluating Learning

Algorithms. IEEE Trans Knowl Data Eng. 2005;17: 299–310.

74.Cohen J. A Coefficient of Agreement for Nominal Scales. Educ Psychol Meas. 1960;20: 37–46.

75.Vieira SM, Kaymak U, Sousa JMC. Cohen's kappa coefficient as a performance measure for feature selection. International Conference on Fuzzy Systems. Barcelona, Spain: IEEE; 2010. pp. 1–8.

76.McHugh M. Interrater reliability: The kappa statistic. Biochem Medica Časopis Hrvat Druš Med Biokem HDMB. 2012;22: 276–82. pmid:23092060

77.Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33: 159–174. pmid:843571

78.Banerjee M, Capozzoli M, McSweeney L, Sinha D. Beyond kappa: A review of interrater agreement measures. Can J Stat. 1999;27: 3–23.

79.Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH. The WEKA Data Mining Software: An Update. SIGKDD Explor Newsl. 2009;11: 10–18.

80.Bravo-Marquez F, Frank E, Pfahringer B, Mohammad SM. AffectiveTweets: a Weka package for analyzing affect in tweets. 2019;20: 1–6.

81.Ptaszynski M, Eronen JKK, Masui F. Learning Deep on Cyberbullying is Always Better Than Brute Force. 2017; 8.

82.Al-Garadi MA, Hussain MR, Khan N, Murtaza G, Nweke HF, Ali I, et al. Predicting Cyberbullying on Social Media in the Big Data Era Using Machine Learning Algorithms: Review of Literature and Open Challenges. IEEE Access. 2019;7: 70701–70718.

83.Sundararaman A, Valady Ramanathan S, Thati R. Novel Approach to Predict Hospital Readmissions Using Feature Selection from Unstructured Data with Class Imbalance. Big Data Res. 2018;13: 65–75.

84.Davis J, Goadrich M. The relationship between Precision-Recall and ROC curves. Proceedings of the 23rd international conference on Machine learning. New York, NY, USA: Association for Computing Machinery; 2006. pp. 233–240.