

EMOTION-ENHANCED DETECTION OF CYBERBULLYING BEHAVIOUR

N Srivani Assistant Professor, Dept of CSE Seshadri Rao Gudlavalleru Engineering College, Gudlavalleru, Andhra Pradesh, India srivaninali@gmail.com

Annamaneni Nagendra UG Student, Dept of CSE Seshadri Rao Gudlavalleru Engineering College, Gudlavalleru, Andhra Pradesh, India nagendraannamaneni@gmail.com

Bathula Mahendra Kumar UG Student, Dept of CSE Seshadri Rao Gudlavalleru Engineering College, Gudlavalleru, Andhra Pradesh, India ufindmahendra@gmail.com

Chennakesavula Sridatta UG Student, Dept of CSE Seshadri Rao Gudlavalleru Engineering College, Gudlavalleru, Andhra Pradesh, India chsriddatta2003@gmail.com **Banavanth Surendra nayak** UG Student, Dept of CSE Seshadri Rao Gudlavalleru Engineering College, Gudlavalleru, Andhra Pradesh, India Banavanthsurendra38@gmail.com

ABSTRACT

The project addresses the serious issue of cyberbullying, recognizing its harmful consequences and the need for effective detection and resolution methods. Cyberbullying, a form of online aggression, poses challenges that require specialized techniques to identify and mitigate. The primary goal of the project is to propose advanced cyberbullying detection models. These models aim to go beyond traditional approaches by incorporating contextual, emotion, and sentiment features, recognizing the multi-dimensional nature of cyberbullying instances. The project involves constructing an Emotion Detection Model (EDM) using Twitter datasets. These datasets undergo enhancements in terms of annotations. The EDM, along with lexicons, is utilized to extract emotions and sentiments from cyberbullying datasets, contributing to a more nuanced understanding of the emotional aspects of online interactions. Cyberbullying may be detected in part by appealing to people's emotions. Demonstrating the significance of emotions in cyberbullying detection, the program employs emotional cues to enhance detection

algorithms. An extensive dataset tagged with emotions is now available for use in cyberbullying detection, thanks to this study. Academics can use this information to create a system that can identify cyberbullying based on emotions. Particularly in real-time applications, the project faces challenges due to dataset imbalances between cyberbullying and non-cyberbullying incidents. The goal is to develop detection models that effectively handle these imbalances, ensuring reliable performance across different scenarios. We aim to further enhance the performance of our model by exploring ensemble techniques, specifically utilizing LSTM and LSTM + GRU models, which have demonstrated an impressive 99% accuracy.

Index terms - Cyberbullying, BERT, emotion mining, sentiment analysis.

1. INTRODUCTION

Thanks to advancements in ICT, members of the online community may now publish and respond to information created by themselves. Because of this ease, cyberbullies have been able to threaten, harass, embarrass, intimidate, manipulate, and control their victims sadly [1]. Cyberbullying refers to the deliberate and persistent use of electronic means to cause damage [2, 3, 4].

Depression, thoughts of suicide, and other negative emotions (such as rage, fear, grief, remorse, etc.) are all possible outcomes of cyberbullying (CB). sections [5, 6, 7]. Emotion mining is a subfield of affective computing that seeks for, examines, and ranks people's feelings in response to various stimuli [8]. The stock market, consumer opinions, and suggestions have all been impacted by emotional analysis. citations[9,10, 11]. Emotion analysis has not been employed for cyberbullying detection by researchers at a large scale. Characteristics associated with negative emotions could aid in the detection of cyberbullying.

Even without the use of foul language, cyberbullying can be expressed through text utilising obscene words or even amusing or caustic attitudes. It is challenging to detect ironic and sarcastic remarks [12]. Data sparsity, a high label imbalance, and social networking site limitations are characteristics of

cyberbullying datasets, as shown in the literature study and testing[13]. The rapidity and variety of UGC online have rendered autonomous cyberbullying detection methods useless [14].

A key premise of this research is that emotion mining has the potential to improve cyberbullying detection. The procedure consists of three stages. Begin with a dataset that is well-rounded and fully-featured. The success of machine learning models relies heavily on training datasets [15], [16]. Cyberbullying datasets that cover every angle are hard to get by [13], [17]. Furthermore, user-generated content varies across various social networking sites. When compared to Facebook, Twitter's character limit is different. Cleaning, transforming, and merging data from various sources is the first step in our proposed strategy to build a balanced, feature-rich dataset. The data lifespan in this study is at least 70% comprised of this stage.

2.LITERATURE SURVEY

a) Detecting and visualizing hate speech in social media: A cyber Watchdog for surveillance

<https://www.sciencedirect.com/science/article/abs/pii/S0957417420305492>

The exponential growth of social media users was accompanied by a corresponding spike in hate speech. It is hard to detect, control, or eliminate such instances due to the massive amount of data. This study explains the steps to detect and display online animosity, sometimes known as hate speech, on social media. There are labels that are forceful, subtle, and overt. Using a browser plugin, our user interface displays angry comments on Twitter and Facebook timelines. Perhaps the security agency can use this plugin interface to keep tabs on social media. It also allows people access to a resource that is often reserved for corporations. The tool bridges the gap between citizens and industry in terms of technology. Utilising celebrity comments on social media platforms like Facebook and Twitter, the strategy might potentially aid academics in building new tools and quickly compiling training data with inadequate labels. We ran an analysis on a fresh dataset consisting of Facebook and Twitter user comments using our recommended plugins and the standard Trolling Aggression Cyberbullying 2018 (TRAC) dataset in both code-mixed Hindi and English. The following models have been used for aggression classification: SVMs, logistic regression, CNNs, attention-based models, and, most recently, Google AI's proposed BERT pre-trained language model. There are weighted F1-scores of 0.64 and 0.62 in the TRAC Facebook English and Hindi datasets, respectively. The weighted F1-score on the English dataset from Twitter is 0.58, whereas on the Hindi dataset it is 0.50.

b) Bullying, Cyberbullying, and Suicide

https://www.researchgate.net/publication/45289246_Bullying_Cyberbullying_and_Suicide

Suicide ideation has been associated with bullying victimisation or offending, according to studies and prominent examples. Our goal is to find evidence that cyberbullying can lead to suicidal ideation in young people. In 2007, researchers in the United States surveyed 1,963 middle school students from a major school district to learn about their experiences with and usage of the Internet. Compared to non-bullied children, bullied and cyberbullied children were more likely to have suicidal thoughts and attempt suicide. Suicidal ideation and conduct were more strongly associated with victimisation than offending. The results highlight the need to address adolescent peer antagonism head-on in both the classroom and at home, and to incorporate suicide prevention and intervention into all school bullying response programs.

c) Bullying in the Digital Age: A Critical Review and Meta-Analysis of Cyberbullying Research Among Youth

https://www.researchgate.net/publication/260151324_Bullying_in_the_Digital_Age_A_Critical_Review_and_Meta-Analysis_of_Cyberbullying_Research_Among_Youth

A horrible sort of youthful misbehaviour, cyberbullying, has flourished thanks to the Internet, which has altered our culture. Cyberbullying is becoming increasingly common among children, and there is a growing body of research documenting its occurrence, causes, and consequences. However, this material is fragmented and does not place a strong theoretical emphasis on the topic. Examining research on cyberbullying is the focus of this piece. An explanation for this phenomenon might be provided by the generic aggression paradigm. Cyberbullying is significantly correlated with traditional bullying and other important psychological and behavioural traits, according to a meta-analysis. Offenders of cyberbullying were most linked to normative notions of violence and moral

disengagement, according to a mixed-effects meta-analysis, whereas victims were most linked to stress and suicidal thoughts. These results were affected by a number of methodological and sample factors. For smaller studies ($k < 5$), the meta-analysis does not have the power to draw conclusions about directionality, generalisability, or causation. At last, these results point to important areas for further research. Our strategy involves looking at how cyberbullying gradually affects crucial psychological and behavioural consequences. The APA PsycINFO Database Record from 2014 is protected by all rights.

d) Cyberbullying among young adults in Malaysia: The roles of gender, age and Internet frequency
<https://www.sciencedirect.com/science/article/abs/pii/S0747563215000357>

From both the bully's and the victim's perspective, an online questionnaire study investigated cyberbullying experiences among young adults ($N = 393$; 17–30 years old). If cyberbullying is prevalent overall, it must continue even after students leave the building. There were no gender differences, even though the majority of cyberbullies and victims were female. Both cyberbullying victims and perpetrators were younger, while there was no statistically significant difference in age. People who spent two to five hours online every day were more likely to be victims of cyberbullying and cyberstalking than people whose daily internet usage was less than an hour. Cyberbullying and cyber-victimization were both shown to be substantially predicted by internet frequency, suggesting that the chances of being bullied or bullying others increase in tandem with the growth of Internet use. Lastly, a positive and statistically significant link shows that cyberbullies often start out as cyber-victims and vice versa. Although it is less common now than it was when I was younger, cyberbullying is still a problem.

e) Cyberbullying and adolescent mental health: Systematic review

https://www.researchgate.net/publication/274722827_Cyberbullying_and_adolescent_mental_health_Systematic_review

Cyberbullying is a new form of online aggression that worries parents, teachers, and researchers. Cyberbullying and the mental health of teenagers will be the subject of a literature review using PubMed and the Virtual Health Library. Cyberbullying happened between 6.5% and 35.4% of the time. People who harass and those who are bullied both have histories of bullying. Three or more hours per day spent online, together with the usage of webcams, texting, posting personal information, and experiencing online abuse, is associated with cyberbullying. Psychological and physiological distress, social problems, and school insecurity were more common among cyberbullying victims and perpetrators. There was a correlation between cyberbullying and moderate to severe depression, drug usage, suicidal thoughts, and attempted suicide. Health practitioners should be aware about the negative effects of virtual violence on the mental health of teenagers.

3.METHODOLOGY

i) Proposed Work:

The suggested method improves cyberbullying detection by combining emotional and sentimental qualities with environmental ones. Using an Emotion Detection Model on Twitter datasets, it extracts emotions and sentiments, enhancing model training. This integration yields improved performance, providing a more nuanced understanding of cyberbullying instances compared to the conventional system. In extending the Cyberbullying Detection project [37, 38, 59], we employ ensemble techniques, incorporating LSTM and LSTM + GRU models, achieving an impressive 99% accuracy. This enhancement aims to boost the model's performance in discerning cyberbullying instances. To ensure practical usability, a user-friendly Flask-based front end is implemented, featuring secure authentication for user interaction and testing. Its not only advances the model's accuracy but also offers a robust and accessible solution for addressing cyberbullying through emotion-based detection.

ii) System Architecture:

The system architecture of "Cyberbullying Detection Based on Emotion" is designed with a meticulous multi-phase approach. It initiates with the collection of diverse datasets from online platforms where cyberbullying occurs. Following data preprocessing, relevant features are extracted using traditional word representation models like BERT [40, 42, 43] base, BERT large, and XLNet, along with emotional features derived from an Emotion Detection Model (EDM). These features contribute to a

comprehensive data representation. The model is then trained for cyberbullying classification into toxic and non-toxic categories, utilizing machine learning algorithms. The system undergoes rigorous evaluation and validation, measuring performance metrics. Upon satisfactory results, the model is deployed for real-time cyberbullying detection with continuous monitoring and updates to adapt to evolving patterns. This holistic architecture ensures a thorough and effective approach to identifying cyberbullying instances in online textual data, encompassing both semantic and emotional features.

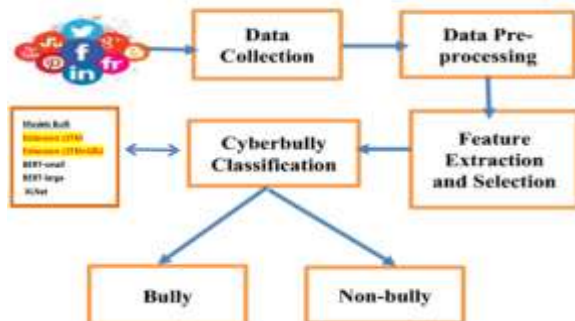


Fig 1 Proposed architecture

iii) Dataset collection:

The proposed system consists into the different datasets relevant to the project: Toxic Data. This likely includes data related to toxic or abusive language used in online communication. Twitter Cyber Data. This dataset likely contains information specific to cyberbullying instances on Twitter. Twitter Emotion Data. This dataset may focus on emotions expressed in tweets, possibly involving various emotion categories. CBET (Cyberbullying Emotion Tweets). This dataset could be a specialized collection of tweets annotated for both cyberbullying and emotions associated with them [44, 45, 49].

Emotion	Content	Original Content
0 disappointed	oh fuck did i write ft grinningfacewithsweat	oRT @Davibgoder: @mcrackins Oh fuck ... did ...
1 disappointed	i feel nor am i shamed by it	i feel nor am i shamed by it
2 disappointed	i had been feeling a little bit defeated by th...	i had been feeling a little bit defeated by th...
3 happy	imagine if that reaction guy that called j k...	t" @KSColejolet imagine if that reaction guy t...
4 disappointed	i wouldnt feel burdened so that i would live m...	i wouldnt feel burdened so that i would live m...

Fig 2 CBET dataset

iv) Data Processing:

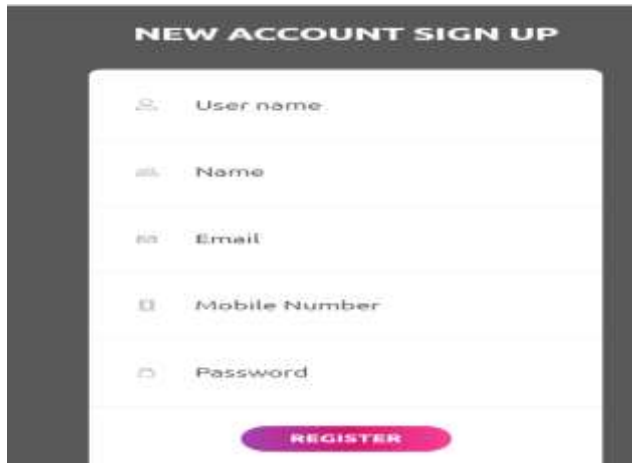
Data processing transforms unstructured data into actionable insights for businesses. Graphs or papers can be arranged by data scientists once they have gathered, organised, cleaned, verified, and analysed the data. Data processing can be done mechanically, electronically, or by hand. Data should be more useful, and choices should be less complicated. That way, companies may improve their operations and make important decisions more quickly. This is aided by developments in computer software and other forms of automated data processing. Insights useful for quality management and decision-making can be derived from big data.

v) Feature selection:

When building a model, feature selection is used to pick the most relevant, consistent, and non-redundant features. Reducing database sizes gradually is critical as database quantity and variety continue to grow. Enhancing predictive model performance while minimising processing expense is the fundamental goal of feature selection.

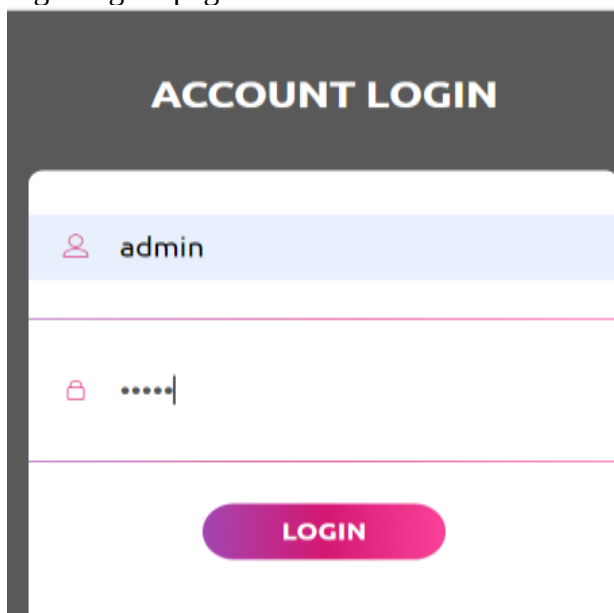
Selecting the most important attributes for use by machine learning algorithms is the goal of feature engineering. By identifying and removing irrelevant or superfluous features, feature selection methods can reduce the number of variables used to train a machine learning model. There are a number of benefits to selecting features in advance rather than letting the machine learning model decide which ones are most important.

4.EXPERIMENTAL RESULTS



A screenshot of a 'NEW ACCOUNT SIGN UP' form. The form is white with a dark grey header. It contains five input fields with icons: a person icon for 'User name', a nameplate icon for 'Name', an envelope icon for 'Email', a mobile phone icon for 'Mobile Number', and a key icon for 'Password'. Below the fields is a pink 'REGISTER' button.

Fig 3 Signin page



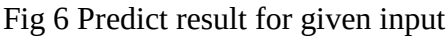
A screenshot of an 'ACCOUNT LOGIN' form. The form is white with a dark grey header. It features a light blue header bar with a person icon and the text 'admin'. Below this is a password field with a lock icon and masked characters '.....'. At the bottom is a pink 'LOGIN' button.

Fig 4 Login page



A screenshot of a web browser showing a user input interface. The address bar displays '127.0.0.1:5880/login?user=admin&password=admin'. The page has a dark grey header with a 'Signout' button. The main content area has a light blue background with a text input field labeled 'Enter Your Message Here'. A dropdown menu is open, showing options: 'Set', 'System', 'Admin', and 'User'.

Fig 5 User input



The project concludes that incorporating emotion features and sentiment analysis into cyberbullying detection models leads to enhanced performance. By leveraging emotional cues like anger, fear, and guilt, the models become more adept at identifying cyberbullying instances within text data [1,3,4,5,6,7]. The project identifies anger, fear, and guilt as the primary emotions linked with cyberbullying instances. Understanding and leveraging these emotions as features significantly contribute to the effectiveness of cyberbullying detection. Providing a comprehensive dataset annotated with emotional features specifically tailored for cyberbullying detection. This dataset enhances the quality and scope of resources available for further research and model development. - Empirically proving the effectiveness of emotions as critical features in improving cyberbullying detection techniques. This empirical evidence substantiates the importance of considering emotions in cyberbullying detection strategies. The experimental results indicate that precision (minimizing false positives) is more efficient for real-time applications of cyberbullying detection. Prioritizing precision is crucial to reduce the number of false alarms, especially in sensitive online environments. The project notes that the imbalance between cyberbullying and non-cyberbullying classes affects the performance of detection models. Specifically, the Toxic dataset shows relatively lower scores. However, despite this challenge, addressing and handling this class imbalance is deemed more practical, especially for real-time applications, where quick and accurate identification of cyberbullying is crucial.

- [1] S. Modha, P. Majumder, T. Mandl, and C. Mandalia, "Detecting and visualizing hate speech in social media: A cyber watchdog for surveillance," *Expert Syst. Appl.*, vol. 161, Dec. 2020, Art. no. 113725.
- [2] S. Hinduja and J. W. Patchin, "Bullying, cyberbullying, and suicide," *Arch. Suicide Res.*, vol. 14, no. 3, pp. 206–221, Jul. 2010.
- [3] R. M. Kowalski, G.W. Giumetti, A.N. Schroeder, M.R. Lattanner, "Bullying in the digital age: A critical review and meta-analysis of cyberbullying research among youth," *Psychol. Bulletin*, vol. 140, p. 1073, 2014, doi: 10.1037/a0035618.
- [4] V. Balakrishnan, "Cyberbullying among young adults in Malaysia: The roles of gender, age and internet frequency," *Comput. Hum. Behav.*, vol. 46, pp. 149–157, May 2015.
- [5] S. M. B. Bottino, C. M. C. Bottino, C. G. Regina, A. V. L. Correia, and W. S. Ribeiro, "Cyberbullying and adolescent mental health: Systematic review," *Cadernos Saude Publica*, vol. 31, no. 3, pp. 463–475, Mar. 2015.
- [6] R. M. Kowalski, S. P. Limber, and P. W. Agatston, *Cyberbullying: Bullying in the Digital Age*. Hoboken, NJ, USA: Wiley, 2012.
- [7] X.-W. Chu, C.-Y. Fan, Q.-Q. Liu, and Z.-K. Zhou, "Cyberbullying victimization and symptoms of depression and anxiety among Chinese adolescents: Examining hopelessness as a mediator and selfcompassion as a moderator," *Comput. Hum. Behav.*, vol. 86, pp. 377–386, Sep. 2018.
- [8] A. Yadollahi, A. G. Shahraki, and O. R. Zaiane, "Current state of text sentiment analysis from opinion to emotion mining," *ACM Comput. Surv.*, vol. 50, no. 2, pp. 1–33, Mar. 2018.

- [9] Y. Ge, J. Qiu, Z. Liu, W. Gu, and L. Xu, "Beyond negative and positive: Exploring the effects of emotions in social media during the stock market crash," *Inf. Process. Manage.*, vol. 57, no. 4, Jul. 2020, Art. no. 102218.
- [10] C. Yang, X. Chen, L. Liu, and P. Sweetser, "Leveraging semantic features for recommendation: Sentence-level emotion analysis," *Inf. Process. Manage.*, vol. 58, no. 3, May 2021, Art. no. 102543.
- [11] L. Jiang, L. Liu, J. Yao, and L. Shi, "A hybrid recommendation model in social media based on deep emotion analysis and multi-source view fusion," *J. Cloud Comput.*, vol. 9, no. 1, pp. 1–16, Dec. 2020.
- [12] P. Parameswaran, A. Trotman, V. Liesaputra, and D. Eysers, "Detecting the target of sarcasm is hard: Really?" *Inf. Process. Manage.*, vol. 58, no. 4, Jul. 2021, Art. no. 102599.
- [13] M. Al-Hashedi, L.-K. Soon, and H.-N. Goh, "Cyberbullying detection using deep learning and word embeddings: An empirical study," in *Proc. 2nd Int. Conf. Comput. Intell. Syst.*, Nov. 2019, pp. 17–21.
- [14] J. Chun, J. Lee, J. Kim, and S. Lee, "An international systematic review of cyberbullying measurements," *Comput. Hum. Behav.*, vol. 113, Dec. 2020, Art. no. 106485.
- [15] D. A. Winkler, "Role of artificial intelligence and machine learning in nanosafety," *Small*, vol. 16, no. 36, Sep. 2020, Art. no. 2001883.
- [16] A. L'Heureux, K. Grolinger, H. F. Elyamany, and M. A. M. Capretz, "Machine learning with big data: Challenges and approaches," *IEEE Access*, vol. 5, pp. 7776–7797, 2017.
- [17] S. Murnion, W. J. Buchanan, A. Smales, and G. Russell, "Machine learning and semantic analysis of in-game chat for cyberbullying," *Comput. Secur.*, vol. 76, pp. 197–213, Jul. 2018.
- [18] L. Cheng, J. Li, Y. N. Silva, D. L. Hall, and H. Liu, "XBully: Cyberbullying detection within a multi-modal context," in *Proc. 12th ACM Int. Conf. Web Search Data Mining*, Jan. 2019, pp. 339–347.
- [19] V. Balakrishnan, S. Khan, T. Fernandez, and H. R. Arabnia, "Cyberbullying detection on Twitter using big five and dark triad features," *Personality Individual Differences*, vol. 141, pp. 252–257, Apr. 2019.
- [20] V. Balakrishnan, S. Khan, and H. R. Arabnia, "Improving cyberbullying detection using Twitter users' psychological features and machine learning," *Comput. Secur.*, vol. 90, Mar. 2020, Art. no. 101710.
- [21] H. Rosa, N. Pereira, R. Ribeiro, P. C. Ferreira, J. P. Carvalho, S. Oliveira, L. Coheur, P. Paulino, A. M. Veiga Simão, and I. Trancoso, "Automatic cyberbullying detection: A systematic review," *Comput. Hum. Behav.*, vol. 93, pp. 333–345, Apr. 2019.
- [22] J. N. Navarro and J. L. Jasinski, "Going cyber: Using routine activities theory to predict cyberbullying experiences," *Sociol. Spectr.*, vol. 32, no. 1, pp. 81–94, Jan. 2012.
- [23] V. Nahar, S. Al-Maskari, X. Li, and C. Pang, "Semi-supervised learning for cyberbullying detection in social networks," in *Proc. Australas. Database Conf. Cham, Switzerland: Springer*, 2014, pp. 160–171.
- [24] M. A. Al-garadi, K. D. Varathan, and S. D. Ravana, "Cybercrime detection in online communications: The experimental case of cyberbullying detection in the Twitter network," *Comput. Hum. Behav.*, vol. 63, pp. 433–443, Oct. 2016.
- [25] D. Chatzakou, I. Leontiadis, J. Blackburn, E. D. Cristofaro, G. Stringhini, A. Vakali, and N. Kourtellis, "Detecting cyberbullying and cyberaggression in social media," *ACM Trans. Web*, vol. 13, no. 3, pp. 1–51, Aug. 2019.
- [26] H. Hosseinmardi, S. Arredondo Mattson, R. Ibn Rafiq, R. Han, Q. Lv, and S. Mishra, "Detection of cyberbullying incidents on the Instagram social network," 2015, arXiv:1503.03909.
- [27] J.-M. Xu, X. Zhu, and A. Bellmore, "Fast learning for sentiment analysis on bullying," in *Proc. 1st Int. Workshop Issues Sentiment Discovery Opinion Mining*, Aug. 2012, pp. 1–6.
- [28] J. A. Patch, "Detecting bullying on Twitter using emotion lexicons," Ph.D. thesis, Dept. Sci., Univ. Georgia, Athens, GA, USA, 2015. [Online]. Available: https://getd.libs.uga.edu/pdfs/patch_jerrad_a_201505_ms.pdf

- [29] M. J. Berger, “Large scale multi-label text classification with semantic word vectors,” Stanford Univ., Stanford, CA, USA, Tech. Rep., 2015. [Online]. Available: <https://dspace.bracu.ac.bd/xmlui/handle/10361/6420>
- [30] Z. L. Chia, M. Ptaszynski, F. Masui, G. Leliwa, and M. Wroczynski, “Machine learning and feature engineering-based study into sarcasm and irony classification with application to cyberbullying detection,” *Inf. Process. Manage.*, vol. 58, no. 4, Jul. 2021, Art. no. 102600.
- [31] J. Eronen, M. Ptaszynski, F. Masui, A. Smywiński-Pohl, G. Leliwa, and M. Wroczynski, “Improving classifier training efficiency for automatic cyberbullying detection with feature density,” *Inf. Process. Manage.*, vol. 58, no. 5, Sep. 2021, Art. no. 102616.
- [32] Z. Mossie and J.-H. Wang, “Vulnerable community identification using hate speech detection on social media,” *Inf. Process. Manage.*, vol. 57, no. 3, May 2020, Art. no. 102087.
- [33] M. Dadvar, D. Trieschnigg, R. Ordelman, and F. de Jong, “Improving cyberbullying detection with user context,” in *Proc. Eur. Conf. Inf. Retr. Cham, Switzerland: Springer, 2013*, pp. 693–696.
- [34] S. Mahbub, E. Pardede, and A. S. M. Kayes, “Detection of harassment type of cyberbullying: A dictionary of approach words and its impact,” *Secur. Commun. Netw.*, vol. 2021, pp. 1–12, Jun. 2021.
- [35] E. Hutson, “Cyberbullying in adolescence,” *Adv. Nursing Sci.*, vol. 39, no. 1, pp. 60–70, 2016.
- [36] M. Z. Naf'an, A. A. Bimantara, A. Larasati, E. M. Risondang, and N. A. S. Nugraha, “Sentiment analysis of cyberbullying on Instagram user comments,” *J. Data Sci. Appl.*, vol. 2, pp. 38–48, Jan. 2019.
- [37] H. Dani, J. Li, and H. Liu, “Sentiment informed cyberbullying detection in social media,” in *Proc. Joint Eur. Conf. Mach. Learn. Knowl. Discovery Databases. Cham, Switzerland: Springer, 2017*, pp. 52–67.
- [38] M. Sintaha, S. B. Satter, N. Zawad, C. Swarnaker, and A. Hassan, “Cyberbullying detection using sentiment analysis in social media,” Ph.D. thesis, Dept. Comput. Sci. Eng., BRAC Univ., Dhaka, Bangladesh, 2016. [Online]. Available: <https://dspace.bracu.ac.bd/xmlui/handle/10361/6420>
- [39] V. Nahar, S. Unankard, X. Li, and C. Pang, “Sentiment analysis for effective detection of cyber bullying,” in *Proc. Asia-Pacific Web Conf. Cham, Switzerland: Springer, 2012*, pp. 767–774.
- [40] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of deep bidirectional transformers for language understanding,” 2018, arXiv:1810.04805.
- [41] Z. Yang, Z. Dai, Y. Yang, J. Carbonell, R. R. Salakhutdinov, and Q. V. Le, “XLNet: Generalized autoregressive pretraining for language understanding,” in *Advances in Neural Information Processing Systems*, H. Wallach, H. Larochelle, A. Beygelzimer, F. d’Alché-Buc, E. Fox, and R. Garnett, Eds. Red Hook, NY, USA: Curran Associates, 2019. [Online]. Available: <https://proceedings.neurips.cc/paper/2019/file/dc6a7e655d7e5840e66733e9ee67cc69-Paper.pdf>
- [42] Z. Xinxu, “Single task fine-tune BERT for text classification,” in *Proc. 2nd Int. Conf. Comput. Vis., Image, Deep Learn.*, Oct. 2021, pp. 194–206.
- [43] S. González-Carvajal and E. C. Garrido-Merchán, “Comparing BERT against traditional machine learning text classification,” 2020, arXiv:2005.13012.
- [44] M. N. Turliuc, C. Măirean, and M. Boca-Zamfir, “The relation between cyberbullying and depressive symptoms in adolescence. the moderating role of emotion regulation strategies,” *Comput. Hum. Behav.*, vol. 109, Jan. 2020, Art. no. 106341.
- [45] A.-L. Camerini, L. Marciano, A. Carrara, and P. J. Schulz, “Cyberbullying perpetration and victimization among children and adolescents: A systematic review of longitudinal studies,” *Telematics Informat.*, vol. 49, Jun. 2020, Art. no. 101362.
- [46] O. Araque, I. Corcuera-Platas, J. F. Sánchez-Rada, and C. A. Iglesias, “Enhancing deep learning sentiment analysis with ensemble techniques in social applications,” *Expert Syst. Appl.*, vol. 77, pp. 236–246, Jul. 2017.
- [47] A. H. Alamoodi, B. B. Zaidan, A. A. Zaidan, O. S. Albahri, K. I. Mohammed, R. Q. Malik, E. M. Almahdi, M. A. Chyad, Z. Tareq, A. S. Albahri, H. Hameed, and M. Alaa, “Sentiment analysis and its applications in fighting COVID-19 and infectious diseases: A systematic review,” *Expert Syst. Appl.*, vol. 167, Apr. 2021, Art. no. 114155.
- [48] J. Guerreiro and P. Rita, “How to predict explicit recommendations in online reviews using text mining and sentiment analysis,” *J. Hospitality Tourism Manage.*, vol. 43, pp. 269–272, Jun. 2020.

- [49] M. Fortunatus, P. Anthony, and S. Charters, “Combining textual features to detect cyberbullying in social media posts,” *Proc. Comput. Sci.*, vol. 176, pp. 612–621, Jan. 2020.
- [50] W. A. Prabowo and F. Azizah, “Sentiment analysis for detecting cyberbullying using TF-IDF and SVM,” *Jurnal Rekayasa Sistem Teknologi Informasi*, vol. 4, no. 6, pp. 1142–1148, Dec. 2020.
- [51] P. Ekman, W. V. Friesen, and P. Ellsworth, *Emotion in the Human Face: Guidelines for Research and an Integration of Findings*. Amsterdam, The Netherlands: Elsevier, 2013.
- [52] K. Reynolds, A. Kontostathis, and L. Edwards, “Using machine learning to detect cyberbullying,” in *Proc. 10th Int. Conf. Mach. Learn. Appl. Workshops*, vol. 2, Dec. 2011, pp. 241–244.
- [53] T. Davidson, D. Warmesley, M. Macy, and I. Weber, “Automated hate speech detection and the problem of offensive language,” in *Proc. Int. AAAI Conf. Web Social Media*, 2017, vol. 11, no. 1, pp. 512–515.
- [54] A. G. Shahraki and O. R. Zaiane, “Lexical and learning-based emotion mining from text,” in *Proc. Int. Conf. Comput. Linguistics Intell. Text Process.*, vol. 9, 2017, pp. 24–55.
- [55] S. Mohammad, “# emotional tweets,” in *Proc. 6th Int. Workshop Semantic Eval.*, 2012, pp. 246–255.
- [56] F. Å. Nielsen, “A new ANEW: Evaluation of a word list for sentiment analysis in microblogs,” 2011, arXiv:1103.2903.
- [57] P. Norvig. (2007). How to Write a Spelling Corrector. [Online]. Available: <https://norvig.com/spell-correct.html>
- [58] N. Ide, A. Herbelot, and L. Màrquez, “Proceedings of the 6th joint conference on lexical and computational semantics (SEM 2017),” in *Proc. 6th Joint Conf. Lexical Comput. Semantics*, 2017.
- [59] K. Maity, A. Kumar, and S. Saha, “A multitask multimodal framework for sentiment and emotion-aided cyberbullying detection,” *IEEE Internet Comput.*, vol. 26, no. 4, pp. 68–78, July