8 REFERENCES

[1] P. Badjatiya, S. Gupta, M. Gupta and V. Varma, “Deep Learning for Hate Speech Detection in Tweets,” in *Proceedings of ACM WWW'17 Companion*, Perth, 2017.

[2] S. Zimmerman, C. Fox and U. Kruschwitz, “Improving Hate Speech Detection with Deep Learning Ensembles,” in *Language Resources and Evaluation Conference (LREC)*, Miyazaki, Japan, 2018.

[3] R. Alshalan and H. Al-Khalifa, “A Deep Learning Approach for Automatic Hate Speech Detection in the Saudi Twittersphere,” *Applied Sciences*, vol. 10, 2020.

[4] U. Naseem, I. Razzak and I. A. Hameed, “Deep Context-Aware Embedding for Abusive and Hate Speech detection on Twitter,” 2019.

[5] I. Amali and S. Jayalal, “Classification of Cyberbullying Sinhala Language Comments on Social Media,” in *MERcon 2020*, Moratuwa, 2020.

[6] H. Caldera, G. Meedin and I. Perera, “Time Series Based Trend Analysis for Hate Speech in Twitter During COVID 19 Pandemic,” in *20th International Conference on Advances in ICT for Emerging Regions (ICTer)*, Colombo, 2020.

[7] D. S. Dias, M. D. Welikala and N. G. J. Dias, “Identifying Racist Social Media Comments in Sinhala Language Using Text Analytics Models with Machine Learning,” in *2018 International Conference on Advances in ICT for Emerging Regions (ICTer)*, 2018.

[8] N. Hettiarachchi, R. Weerasinghe and R. Pushpanda, “Detecting Hate Speech in Social Media Articles in Romanized Sinhala,” in *20th International Conference on Advances in ICT for Emerging Regions (ICTer)*, 2020.

[9] H. Sandaruwan, S. Lorensuhewa and M. Kalyani, “Sinhala Hate Speech Detection in Social Media using Text Mining and Machine learning,” in *19th International Conference on Advances in ICT for Emerging Regions (ICTer)*, 2019.

[10] Z. Zhang, D. Robinson and J. Tepper, “Hate Speech Detection Using a Convolution-LSTM Based Deep Neural Network,” 2017.

[11] S. MacAvaney, H.-R. Yao, E. Yang,, K. Russell, N. Goharian and O. Frieder, “Hate speech detection: Challenges and solutions,” in *PLOS ONE*, 2019.

[12] Z. Zuping, N. D. Gitari, H. Damien and J. Long, “A Lexicon-based Approach for Hate Speech Detection,” in *International Journal of Multimedia and Ubiquitous Engineering*, 2015.

[13] R. E and W. J, “Learning extraction patterns for subjective expressions,” in *Empirical Methods in Natural Language Processing (EMNLP).*, 2003.

[14] W. J and R. E, “Creating subjective and objective sentence classifiers from

unannotated texts," in *6th International Conference On Intelligent Text Processing and Computational Linguistics*, Mexico, 2005.

[15] E. A and S. F, "SentiWordNet: A Publicly Available Lexical Resource for Opinion Mining," in *5th International Conference on Language Resources and Evaluation*, Genoa, 2006.

[16] A. H. Razavi, D. Inkpen, S. Uritsky and S. Matwin, "Offensive Language Detection Using Multi-level," in *Advances in Artificial Intelligence*, 2010.

[17] I. Witten, E. Frank and J. Gray, Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations, 2008.

[18] M. Hall and E. Frank, "Combining Naive Bayes and Decision Tables," in *FLAIRS*, 2008.

[19] Z. Waseem, T. Davidson, D. Warmsley and I. Weber, "Understanding Abuse: A Typology of Abusive Language Detection Subtasks," in *Association for Computational Linguistics*, Vancouver, BC, Canada, 2017.

[20] "Google Bad Words List," [Online]. Available: https://www.freewebheaders.com/full-list-of-bad-words-banned-by-google/.

[21] C. Nobata, J. Tetreault, A. Thomas, Y. Mehdad and Y. Chang, "Abusive Language Detection in Online User Content," in *International World Wide Web Conference Committee*, 2016.

[22] Y. Lee, S. Yoon and K. Jung, Comparative Studies of Detecting Abusive Language on Twitter, Belgium, 2018.

[23] T. Davidson, D. Warmsley, M. Macy and I. Weber, "Automated Hate Speech Detection and the Problem of Offensive Language," in *Proceedings of the International AAAI Conference on Web and Social Media*, 2017.

[24] L. Gao and R. Huang, "Detecting Online Hate Speech Using Context Aware Models," 2018.

[25] G. K. Pitsilis, H. Ramampiaro and H. Langseth, "Effective hate-speech detection in Twitter data using recurrent neural networks," in *Applied Intelligence*, 2018.

[26] K. Steimel,, D. Dakota,, Y. Chen, and S. Kubler, "Investigating Multilingual Abusive Language Detection: A Cautionary Tale," in *International Conference on Recent Advances in Natural Language Processing*, Varna, Bulgaria, 2019.

[27] S. S. Aluru, B. Mathew, P. Saha1, and A. Mukherjee, "Deep Learning Models for Multilingual Hate Speech Detection," 2020.

[28] A. Arango, J. Pérez and B. Poblete, "Hate Speech Detection is Not as Easy as You May Think: A Closer Look at Model Validation," in *Association for Computing Machinery*, New York, 2019.

[29] P. Mathur, R. R. Shah, R. Sawhney and D. Mahata, "Detecting Offensive

Tweets in Hindi-English Code-Switched Language," in *Sixth International Workshop on Natural Language Processing for Social Media*, Melbourne,, 2008.

[30] V. K. Jha, H. P, V. P. N, V. Vijayana and P. P, "DHOT-Repository and Classification of Offensive Tweets in the Hindi Language," *Procedia Computer Science*, vol. 171, pp. 2324-2333, 2020.

[31] S. Kamble and A. Joshi, "Hate Speech Detection from Code-mixed Hindi-English Tweets Using Deep Learning Models," 2018.

[32] A. Alakrot, L. Murray and N. S. Nikolov, "Towards Accurate Detection of Offensive Language in Online," in *4th International Conference on Arabic Computational Linguistics*, Dubai, 2018.

[33] S. T. Luu, H. P. Nguyen, K. V. Nguyen and N. L.-T. Nguyen, "Comparison Between Traditional Machine Learning Models And Neural Network Models For Vietnamese Hate Speech Detection," in *International Conference on Computing and Communication Technologies (RIVF)*, Vietnam, 2020.

[34] V. Santucci, S. Spina, A. Milani, G. Biondi and G. D. Bari, "Detecting Hate Speech for Italian Language in Social Media," in *EVALITA*, Torino, Italy, 2018.

[35] N. Romim, M. Ahmed, H. Talukder and M. S. Islam, "Hate Speech detection in the Bengali language: A dataset and its baseline evaluation," 2020.

[36] "Tools and resources of Natural Language Processing Center at University of Moratuwa," University of Moratuwa, 2020. [Online]. Available: https://uom.lk/nlp/tools. [Accessed 27 02 2021].