IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

Impact Factor: 6.252

Volume 2, Issue 1, June 2022

A Deep Learning Approach for Robust Detection of Bots in Twitter Using Transformers Model

Mr. K. Pazhanivel[1], Ajai Kumar. B[2], Mageshwaran. M[3], Dhivakar. K[4]

Assistant Professor, Department of Science and Computer Engineering[1]

Students, Department of Science and Computer Engineering[2,3,4]

Anjalai Ammal Mahalingam Engineering College, Kovilvenni, Tiruvarur, Tamil Nadu, India

Abstract: *The volume of audio visual content produced on social networks has increased tremendously in recent decades, and this information is quickly spread and consumed by a large number of people. The disruption of false news sources and bot accounts for disseminating fake news is a possibility in this scenario. Applied research has been supported by promotional information as well as sensitive stuff over the network. Artificial Intelligence will be used to automatically assess the trustworthiness of social media accounts (AI). In this research, we describe a multilingual strategy to using Deep Learning to solve the bot identification problem on Twitter. End-users can utilise machine learning (ML) methodologies to assess the trustworthiness of a Twitter account. To achieve so, a number of tests were carried out using cutting-edge Multilingual Language Models. Construct an encoding of the user account's text-based features, which is then concatenated with the rest of the metadata to build a potential input vector on top of a Bot-DenseNet Dense Network. As a result, this article evaluates the language constraint from prior experiments where the encoding of the language was limited. Only the metadata information or the metadata information along with some other information was examined by the user account. properties of fundamental semantic text The Bot-DenseNet also generates a low-dimensional representation of the data. Within the Information Retrieval (IR) framework, a user account can be utilised for any application.*

Keywords: Bot Detector, Deep Learning, Feature Representation, Language Models, Misinformation Detection, Social Media Mining, Transfer Learning, Transformers

REFERENCES

[1]. Ž. Agić and I. Vulić, ‘‘JW300: A wide-coverage parallel corpus for lowresource languages,’’ in Proc. 57th Annu. Meeting Assoc. Comput. Linguistics. Florence, Italy: Association for Computational Linguistics, Jul. 2019, pp. 3204–3210.

[2]. A. Akbik, T. Bergmann, D. Blythe, K. Rasul, S. Schweter, and A. Vollgraf, ‘‘FLAIR: An easy-to-use framework for state-of-the-art NLP,’’ in Proc. Conf. North Amer. Chapter Assoc. Comput. Linguistics, Demonstrations, 2019, pp. 54–59.

[3]. A. Akbik, D. Blythe, and R. Vollgraf, ‘‘Contextual string embeddings for sequence labeling,’’ in Proc. 27th Int. Conf. Comput. Linguistics, 2018, pp. 1638–1649.

[4]. A. S. M. Alharbi and E. de Doncker, ‘‘Twitter sentiment analysis with a deep neural network: An enhanced approach using user behavioral information,’’ Cogn. Syst. Res., vol. 54, pp. 50–61, May 2019.

[5]. N. R. Aljohani, A. Fayoumi, and S.-U. Hassan, ‘‘Bot prediction on social networks of Twitter in altmetrics using deep graph convolutional networks,’’ Soft Comput., vol. 24, pp. 11109–11120, Jan. 2020.

[6]. M. Arora and V. Kansal, ‘‘Character level embedding with deep convolutional neural network for text normalization of unstructured data for Twitter sentiment analysis,’’ Social Netw. Anal. Mining, vol. 9, no. 1, p. 12, Dec. 2019.

[7]. A. Balestrucci, R. De Nicola, O. Inverso, and C. Trubiani, ‘‘Identification of credulous users on Twitter,’’ in Proc. 34th ACM/SIGAPP Symp. Appl. Comput., Apr. 2019, pp. 2096–2103.

[8]. A. Bhoi and S. Joshi, ‘‘Various approaches to aspect-based sentiment analysis,’’ 2018, arXiv:1805.01984. [Online]. Available: http://arxiv. org/abs/1805.01984

[9]. Z. Chu, S. Gianvecchio, H. Wang, and S. Jajodia, ‘‘Detecting automation of Twitter accounts: Are you a human, bot, or cyborg?'' IEEE Trans. Depend. Sec. Comput., vol. 9, no. 6, pp. 811–824, Nov./Dec. 2012.

[10]. T. Cooijmans, N. Ballas, C. Laurent, Ç. Gülçehre, and A. Courville, ''Recurrent batch normalization,'' 2016, arXiv:1603.09025. [Online]. Available: http://arxiv.org/abs/1603.09025

[11]. S. Cresci, R. Di Pietro, M. Petrocchi, A. Spognardi, and M. Tesconi, ''The paradigm-shift of social spambots: Evidence, theories, and tools for the arms race,'' in Proc. 26th Int. Conf. World Wide Web Companion, 2017, pp. 963–972.

[12]. C. A. Davis, O. Varol, E. Ferrara, A. Flammini, and F. Menczer, ''BotOrNot: A system to evaluate social bots,'' in Proc. 25th Int. Conf. Companion World Wide, 2016, pp. 273–274.

[13]. A. Davoudi, A. Z. Klein, A. Sarker, and G. Gonzalez-Hernandez, ''Towards automatic bot detection in twitter for health-related tasks,'' AMIA Summits Transl. Sci. Proc., vol. 2020, p. 136, May 2020.

[14]. J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, ''BERT: Pre-training of deep bidirectional transformers for language understanding,'' 2018, arXiv:1810.04805. [Online]. Available: http://arxiv.org/abs/1810.04805

[15]. J. Diesner, E. Ferrari, and G. Xu, in Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining. Sydney, NSW, Australia: ACM, Aug. 2017. [Online]. Available: https://dblp.org/rec/bib/conf/asunam/2017, doi: 10.1145/3110025.

[16]. C. D. Santos and M. Gatti, ''Deep convolutional neural networks for sentiment analysis of short texts,'' in Proc. 25th Int. Conf. Comput. Linguistics (COLING), 2014, pp. 69–78.

[17]. L. Floridi and M. Chiriatti, ''GPT-3: Its nature, scope, limits, and consequences,'' Minds Mach., vol. 30, pp. 681–694, Nov. 2020.

[18]. R. Gao, F. Liu, J. Zhang, B. Han, T. Liu, G. Niu, and M. Sugiyama, ''Maximum mean discrepancy is aware of adversarial attacks,'' 2020, arXiv:2010.11415. [Online]. Available: http://arxiv.org/abs/2010.11415

[19]. X. He, X. Du, X. Wang, F. Tian, J. Tang, and T.-S. Chua, ''Outer productbased neural collaborative filtering,'' 2018, arXiv:1808.03912. [Online]. Available: http://arxiv.org/abs/1808.03912

[20]. J. Im, E. Chandrasekharan, J. Sargent, P. Lighthammer, T. Denby, A. Bhargava, L. Hemphill, D. Jurgens, and E. Gilbert, ''Still out there: Modeling and identifying Russian troll accounts on Twitter,'' in Proc. 12th ACM Conf. Web Sci., Jul. 2020, pp. 1–10.

[21]. S. Ioffe and C. Szegedy, ''Batch normalization: Accelerating deep network training by reducing internal covariate shift,'' 2015, arXiv:1502.03167. [Online]. Available: http://arxiv.org/abs/1502.03167

[22]. J. Knauth, ''Language-agnostic Twitter-bot detection,'' in Proc. Int. Conf. Recent Adv. Natural Lang. Process. (RANLP), 2019, pp. 550–558.

[23]. Z. Lin, S. Mu, F. Huang, K. A. Mateen, M. Wang, W. Gao, and J. Jia, ''A unified matrix-based convolutional neural network for finegrained image classification of wheat leaf diseases,'' IEEE Access, vol. 7, pp. 11570–11590, 2019.

[24]. F. Liu, W. Xu, J. Lu, G. Zhang, A. Gretton, and D. J. Sutherland, ''Learning deep Kernels for non-parametric two-sample tests,'' 2020, arXiv:2002.09116. [Online]. Available: http://arxiv.org/abs/2002.09116

[25]. Y. Liu, P. Dmitriev, Y. Huang, A. Brooks, and L. Dong, ''An evaluation of transfer learning for classifying sales engagement emails at large scale,'' 2019, arXiv:1905.01971. [Online]. Available: http://arxiv.org/abs/1905.01971

[26]. P. Lynn, ''The advantage and disadvantage of implicitly stratified sampling,'' Methods, Data, Analyses, vol. 13, no. 2, p. 14, 2019.

[27]. M. Mazza, S. Cresci, M. Avvenuti, W. Quattrociocchi, and M. Tesconi, ''RTbust: Exploiting temporal patterns for botnet detection on Twitter,'' in Proc. 10th ACM Conf. Web Sci., 2019, pp. 183–192.

[28]. A. Minnich, N. Chavoshi, D. Koutra, and A. Mueen, ''BotWalk: Efficient adaptive exploration of Twitter bot networks,'' in Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining, Jul. 2017, pp. 467–474.

[29]. U. Naseem, I. Razzak, K. Musial, and M. Imran, ''Transformer based deep intelligent contextual embedding for Twitter sentiment analysis,'' Future Gener. Comput. Syst., vol. 113, pp. 58–69, Dec. 2020.

[30]. M. Orliński and N. Jankowski, ''Fast t-SNE algorithm with forest of balanced LSH trees and hybrid computation of repulsive forces,'' Knowl.- Based Syst., vol. 206, Oct. 2020, Art. no. 106318.

[31]. J. Pizarro, ''Using N-grams to detect bots on Twitter,'' in Proc. CLEF, Working Notes, 2019, pp. 1–10.

[32]. M. Polignano, M. G. de Pinto, P. Lops, and G. Semeraro, ''Identification of bot accounts in Twitter using 2D CNNs on user-generated contents,'' in Proc. CLEF, Working Notes, 2019, pp. 1–11.

[33]. C. Raffel, N. Shazeer, A. Roberts, K. Lee, S. Narang, M. Matena, Y. Zhou, W. Li, and P. J. Liu, ''Exploring the limits of transfer learning with a unified text-to-text transformer,'' J. Mach. Learn. Res., vol. 21, no. 140, pp. 1–67, 2020.

[34]. J. Rodríguez-Ruiz, J. I. Mata-Sánchez, R. Monroy, O. Loyola-González, and A. López-Cuevas, ''A one-class classification approach for bot detection on Twitter,'' Comput. Secur., vol. 91, Apr. 2020, Art. no. 101715.

[35]. K. Shuang, H. Guo, Z. Zhang, J. Loo, and S. Su, ''A word-building method based on neural network for text classification,'' J. Exp. Theor. Artif. Intell., vol. 31, no. 3, pp. 455–474, May 2019.

[36]. N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, ''Dropout: A simple way to prevent neural networks from overfitting,'' J. Mach. Learn. Res., vol. 15, no. 1, pp. 1929–1958, 2014.

[37]. D. Stojanovski, G. Strezoski, G. Madjarov, and I. Dimitrovski, ''Twitter sentiment analysis using deep convolutional neural network,'' in Proc. Int. Conf. Hybrid Artif. Intell. Syst. Springer, 2015, pp. 726–737. [Online]. Available: https://scholar.googleusercontent. com/scholar.bib?q=info:HnIU7VyTzLUJ:scholar. google.com/& output= citation&scisdr=CgXc4k0kELTt-pJoVlM:AAGBfm0AAAAAYGxtTlO0 qf0SoEojztY ZqYNU1uzAmqAp&sci sig=AAGBfm0AAAAAYGxtTlfX vsJzQ3eCFjVQwVDi0pipTQma&scisf=4&ct=citation&cd=-1&hl=es

[38]. O. Varol, E. Ferrara, C. A. Davis, F. Menczer, and A. Flammini, ''Online human-bot interactions: Detection, estimation, and characterization,'' 2017, arXiv:1703.03107. [Online]. Available: http://arxiv. org/abs/1703.0310

[39]. A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, ''Attention is all you need,'' in Proc. Adv. Neural Inf. Process. Syst., 2017, pp. 5998–6008.

[40]. I. Vogel and P. Jiang, ''Bot and gender identification in Twitter using word and character N-grams,'' in Proc. CLEF, Working Notes, 2019, pp. 1–9.

[41]. B. Wang and C.-C. J. Kuo, ''SBERT-WK: A sentence embedding method by dissecting bert-based word models,'' 2020, arXiv:2002.06652. [Online]. Available: http://arxiv.org/abs/2002.06652

[42]. L Wang, J. Niu, and S. Yu, ''SentiDiff: Combining textual information and sentiment diffusion patterns for Twitter sentiment analysis,'' IEEE Trans. Knowl. Data Eng., vol. 32, no. 10, pp. 2026–2039, Oct. 2020.

[43]. T. Wolf et al., ''HuggingFace's transformers: State-of-the-art natural language processing,'' 2019, arXiv:1910.03771. [Online]. Available: http://arxiv.org/abs/1910.03771

[44]. K. Yang, O. Varol, C. A. Davis, E. Ferrara, A. Flammini, and F. Menczer, ''Arming the public with artificial intelligence to counter social bots,'' Hum. Behav. Emerg. Technol., vol. 1, no. 1, pp. 48–61, Jan. 2019.

[45]. K.-C. Yang, O. Varol, P.-M. Hui, and F. Menczer, ''Scalable and generalizable social bot detection through data selection,'' in Proc. AAAI Conf. Artif. Intell., vol. 34, 2020, pp. 1096–1103.

[46]. C. Zhang, S. Bengio, M. Hardt, B. Recht, and O. Vinyals, ''Understanding deep learning requires rethinking generalization,'' 2016, arXiv:1611.03530. [Online]. Available: http://arxiv.org/abs/1611.03530

[47]. S. Zhang, X. Xu, Y. Pang, and J. Han, ''Multi-layer attention based CNN for target-dependent sentiment classification,'' Neural Process. Lett., vol. 51, no. 3, pp. 2089–2103, Jun. 2020.

[48]. J. Zhu, C. Huang, M. Yang, and G. P. Cheong Fung, ''Context-based prediction for road traffic state using trajectory pattern mining and recurrent convolutional neural networks,'' Inf. Sci., vol. 473, pp. 190–201, Jan. 2019