IJARSCT

ISSN (Online) 2581-9429

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

IJARSCT
Impact Factor: 6.252

Volume 2, Issue 5, May 2022

Analysis of Different Machine Learning Algorithms Used for Spam E-mail Detection

Shubham Zanzad[1], Devansh Thard[2], Tushar Jarare[3], Santosh Shinde[4], Prof. B. S. Gayal[5]

Students, Dept. of Information Technology Engineering[1,2,3,4]
Guide, Dept. of Information Technology Engineering[5]
Sinhgad Academy of Engineering, Pune Maharashtra, India

Abstract: *From business to education, email is now used in almost every industry. Subcategories of email exist, such as ham and spam. Unsolicited email, also known as spam or junk email, is a sort of email that can be used to harm consumers by wasting their time, using up their computer resources, and collecting sensitive information. Every day, the amount of spam sent out increases alarmingly. For email and IoT service providers, spam detection and filtering have suddenly become substantial and pervasive concerns. Email filtration is one of the most essential and well-known advanced spam detection and prevention techniques. Many machine learning and deep learning algorithms have been used for this purpose, including Naive Bayes, decision trees, neural networks, and random forests. This article divides utility research approaches into applicable classifications based on machine learning tactics used in texting systems. The accuracy, precision, recall, and other performance characteristics of these approaches are all well assessed. Finally, broad ideas and prospective study directions are provided.*

Keywords: SVM, Decision Tree, K-Nearest Neighbor, Naïve Bayes, Boosting Algorithm

REFERENCES

[1]. Chao Chen, Yu Wang, Jun Zhang, Yang Xiang, Wanlei Zhou, Statistical Features-Based Real-Time Detection of Drifted Twitter Spam, IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, VOL. 12, NO. 4, APRIL 2017.

[2]. L. Breiman, Random forests, Mach. Learn., vol. 45, no. 1, pp. 5-32, 2001.

[3]. C. Grier, K. Thomas, V. Paxson, and M. Zhang, @spam: The underground on 140 characters or less, in Proc. 17th ACM Conf. Comput. Commun. Security, 2010, pp. 27-37.

[4]. H. Kwak, C. Lee, H. Park, and S. Moon, What is twitter, a social network or a news media? in Proc. 19th Int. Conf. World Wide Web, 2010, pp. 591-600.

[5]. K. Lee, J. Caverlee, and S. Webb, Uncovering social spammers: Social honeypots + machine learning, in Proc. 33rd Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.,2010, pp. 435-442.

[6]. J. Oliver, P. Pajares, C. Ke, C. Chen, and Y. Xiang, An in-depth analysis of abuse on twitter, Trend Micro, Irving, TX, USA, Tech. Rep., Sep. 2014.

[7]. Song, S. Lee, and J. Kim, Spam ltering in twitter using sender-receiver relationship, in Proc. 14th Int. Conf. Recent Adv. Intrusion Detection, 2011, pp. 301-317.

[8]. K. Thomas, C. Grier, D. Song, and V. Paxson, Suspended accounts in retrospect: An analysis of twitter spam, in Proc. ACM SIGCOMM Conf. Internet Meas. Cof., 2011, pp. 243-258.

[9]. C. Yang, R. Harkreader, and G. Gu,Empirical evaluation and new design for fighting evolving twitter spammers, IEEE Trans. Inf. Forensics Security, vol. 8, no. 8, pp. 1280-1293, Aug. 2013.

[10]. S. Yardi, D. Romero, G. Schoenebeck, and D. Boyd, Detecting spam in a twitter network, First Monday, vol. 15, nos. 1-4, pp. 1-13, Jan. 2010.

DOI: 10.48175/IJARSCT-4097