IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 3, November 2022
Impact Factor: 6.252

Smart Spam Comment Detection Using Machine Learning

Nauman Zari[1], Sanket Potghan[2], Sweety Kale[3], Prof. S. D. Babar[4]
Students, Department of Information Technology[1,2,3]
Professor, Department of Computer Engineering[4]
Sinhgad Institute of Technology, Lonavala, Maharashtra, India

Abstract: *The profitability promoted by Google in its well-known video distribution platform YouTube has attracted an increasing number of users. However, such success has also attracted large number of malicious users, which aim to self-promote their videos or circulate viruses and malware. As we know that YouTube offers limited tools for comment moderation, so spam increases very rapidly and that's why comment section of the owners is disabled. It is very difficult to established classification methods for automatic spam filtering since the messages are very short and often widespread with slangs, symbols, and abbreviations. In this paper, we have evaluated several top-performance classification techniques for detecting and analysing spam comments. The statistical analysis of results indicates that, with 99.9% of confidence level, decision trees, logistic regression, Bernoulli Naive Bayes, random forests, linear and Gaussian SVMs are statistically equivalent in maximum rate . Therefore, it is very important to find a way to detect these comments on videos and report them before they are viewed by innocent users.*

Keywords: Spam Comment Detection, Machine Learning, Naïve Bayes, Linear Regression, SVM

REFERENCES

[1]. N. M. Samsudin, C. F. B. MohdFoozy, N. Alias, P.Shamala, N. F. Othaman and W. I. S. Wan Din, "Youtube spam detection framework using Naïve Bayes and logistic regression", Indonesian J. Electr. Eng. Comput. Sci., vol. 14, no. 3, pp. 1508, Jun. 2019.
[2]. K. S. Adewole, N. B. Anuar, A. Kamsin, K. D. Varathan and S. A. Razak, "Malicious accounts: Dark of the social networks", Elsevier, 2017, pp. 41-67.
[3]. R. K. Das, S. S. Dash, K. Das and M. Panda, "Detection of spam in Youtube comments using different classifiers", Advanced Computing and Intelligent Engineering, pp. 201-214, 2020.
[4]. S. Garg Makkar, N. Kumar, M. S. Hossain, A. Ghoneim and M. Alrashoud, "An Efficient Spam Detection Technique Using Machine Learning", IEEE Transactions on Industrial Informatics, vol. 17, no. 2, pp.903-912, Feb. 2021.
[5]. C. Chen, Y. Wang, J. Zhang, Y. Xiang, W. Zhou and G. Min, "Statistical Features- Based Real- Time Detection of Drifted Twitter Spam", IEEE Transactions, April 2017, pp.914-925.
[6]. M. Verma, Divya, S. Sofat, "Techniques to Detect Spammers in Twitter – A Survey", International Journal of Computer Applications, January 2014, Vol. 85, No. 10, pp. 27-32.
[7]. A. Gupta and R. Kaushal, "Improving Spam Detection in Online Social Networks", IEEE, 2015

BIOGRAPHY

- Nauman Zari-An Undergraduate Scholar pursuing Bachelors of Engineering in Computer from Sinhgad Institute of Technology. He is working under the guidance of Prof. S.D. Babar.
- Sanket Potghan - An Undergraduate Scholar pursuing Bachelors of Engineering in Computer from Sinhgad Institute of Technology. He is working under the guidance of Prof. S.D. Babar.
- Sweety Kale- An Undergraduate Scholar pursuing Bachelors of Engineering in Computer from Sinhgad Institute of Technology. He is working under the guidance of Prof. S.D. Babar.