TEST Engineering & Management

January-February 2020
ISSN: 0193-4120 Page No. 10766 - 10768

A Framework for Detecting Spam Reviews in Online Social Media

[1]J. Venkata Krishnakanth, [2]J Sybi Cynthia

[1]UG Scholar, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, India
[2]Assistant Professor, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, India
jvkk1999@gmail.com, cynthia.sybi@gmail.com

Article Info
Volume 82
Page Number: 10766 – 10768
Publication Issue:
January-February 2020

Article History
Article Received*: 18 May 2019*
Revised*: 14 July 2019*
Accepted*: 22 December 2019*
Publication*: 19 February 2020*

Abstract

Now a days, reviews and feedback on Social media play a vital role in the aspect of shopping. Any person can drop their review or feedback in the reviews column and this provides a golden opportunity for the spammers. In this paper, we propose an efficient framework that identifies the spammers (spam reviews) by utilizing the spam features for modelling review datasets as Heterogenous Information Network to map spam detection procedure into a classification problem. Different datasets were collected from various sites like Amazon and Yelp and various experiments were conducted on them by utilizing the importance of spam features to obtain the best results.

Keywords: *Social Media, Spammer, Reviews and Heterogenous Information Network*

1. Introduction

Social Media plays an influential role for customers in selecting their products and services. In addition to that producers started to depend on social media in their advertising campaigns. In olden days people are mostly dependent on written reviews but they are independent of those reviews and decision making process became simpler. It became so only with the help of reviews and feedback available on social media.

Both positive and negative reviews are available on social media and they may encourage or discourage customers in the selection process of their products and services. So, these reviews or feedback create impact on customers potentially. With this golden opportunity, Spammers provide fake reviews and can mislead customer`s opinion and then multiplied by the sharing function of social media over the web.

The proposed framework is totally dependent on network based approach and model reviews as Heterogenous Information Networks. Then a new weighing method for spam features is proposed and is used to determine the importance of each feature in identifying spam reviews from normal reviews. The earlier proposed functions are not accurate but the proposed one obtains accuracy as a built-in function.

2. Literature Review

The work of *J. Donfroet al*. [1] described how internet made the things very simple and why everyone likes online shopping. They also described the number of possibilities for a fake products and reviews. So, some different classification techniques were proposed to develop a framework that classifies different type of datasets.

The work of *M. Ott, C. Cardie, and J. T. Hancock et al.* [2] used different methods of designing, detecting and classifying the different type of datasets and tried to implement the same techniques on different datasets collected from different websites of social media. Moreover, this work provides a clear overview on the behavior of a spammer.

Various reviewed, behavioral and other features were implemented in *M. Ott, Y. Choi, C. Cardie, and J. T. Hancock et al.* [3] and evaluates the both performance and efficiency of those features and provides a basic overview on basic outcomes of different features and the work of *] M. Ott, Y. Choi, C. Cardie, and J. T. Hancock*

January-February 2020
ISSN: 0193-4120 Page No. 10766 - 10768

et al. [2] helped a lot in performing those different detection techniques.

The authors, *Ch. Xu and J. Zhanget al.* [4] used the datasets of different type of written reviews and are converted into different type of pattens by using different type of regression techniques. Moreover, the work summarizes how negative reviews can potentially impact market and economic status of an organization.

Different type of datasets are available and the authors, *N. Jindal and B. Liu et al.* [5] describe the procedure involved in modelling several review datasets collected from different websites of social media and also proposed a framework and algorithm to detect spam users. They also tried to implement them on datasets collected.

Supervised and unsupervised learning can be helpful in classification techniques. The authors, *F. Li, M. Huang, Y. Yang, and X. Zhu. et al.* [6] tries to implement them on datasets collected in [4] and in addition to those learning techniques it also makes use of the algorithm proposed in [5] just to increase the efficiency and performance of the framework.

The authors, *G. Fei, A. Mukherjee, B. Liu, M. Hsu, M. Castellanos, and R. Ghosh et al.* [7] clearly define the concept of metapath, review-user and behavioral-linguistic features. Their work also includes the process of calculating the weights of datasets collected from different datasets collected from different websites of social media using both supervised and unsupervised learning approaches.

In addition to those learning techniques, the authors, *A. j. Minnich, N. Chavoshi, A. Mueen, S. Luan, and M. Faloutsos et al.* [8] proposed a new graph method was used to label the reviews based on their performance and efficiency. They noticed the effect and describe the noticeable effect in determining the weights of most of the functions.

B. Viswanath, M. Ahmad Bashir, M. Crovella, S. Guah, K. P. Gummadi, B. Krishnamurthy, and A. Mislove and et al.[9]proposed an approach to calculate the weights of real world labelled datasets collected from Yelp and Amazon websites. On the other hand, they also describes how we can determine their performance in terms of their involvement in connecting spam reviews.

H. Li, Z. Chen, B. Liu, X. Wei, and J. Shao and et al. [10] provides a complete overview on the concepts of Heterogenous Information Networks, Support vector Machine and many others. They even describe how the problem of spam detection can be considered as a new research line in such networks field.

L. Akoglu, R. Chandy, and C. Faloutsos and et al. [11] clearly shows how metapath can increase the efficiency and also describe in detail how it can be used for connected through features of a spammer. In addition to the concept of metapath used in finding the community, it also provides basic information regarding the patterns of different datasets.

The authors, *A. Mukherjee, A. Kumar, B. Liu, J. Wang, M. Hsu, M. Castellanos, and R. Ghosh and et al.*[12] implements those behavioral and linguistic features and describes that linguistic ones work very bad compared to behavioral features. They also uses a new graph based methodology to label the reviews based on rank based labelling approach.

3. Conclusion

This work describes spam detection framework in particular it uses the concept of a meta path plan and another graph based strategy or framework in classifying or analyzing reviews depending on a rank-based naming methodology. The execution of the proposed framework is assessed by utilizing review datasets. The perceptions used in this work demonstrate that discovered weights by utilizing this metapath plan can be exceptionally powerful in recognizing spam reviews and prompts a superior execution. Moreover, it found that only a prepared set of web Spam features will figure the importance of each part and it yields better execution. The outcome also shows the impact of that utilizing numerous supervisions in terms of efficiency, just like the semi administered strategy. Such impact can be mostly decided on their weights as similarly in numerous datasets.

References

[1] J. Donfro, A whopping 20 % of yelp reviews are fake. http://www.businessinsider.com/20-percent-of-yelp-reviews-fake-2013-9. Accessed: 2015-07-30.

[2] M. Ott, C. Cardie, and J. T. Hancock. Estimating the prevalence of deception in online review communities. In ACM WWW, 2012.

[3] M. Ott, Y. Choi, C. Cardie, and J. T. Hancock. Finding deceptive opinion spam by any stretch of the imagination. In ACL, 2011.

[4] Ch. Xu and J. Zhang. Combating product review spam campaigns via multiple heterogeneous pairwise features. In SIAM International Conference on Data Mining, 2014.

[5] N. Jindal and B. Liu. Opinion spam and analysis. In WSDM, 2008.

[6] F. Li, M. Huang, Y. Yang, and X. Zhu. Learning to identify review spam. Proceedings of the 22nd International Joint Conference on Artificial Intelligence; IJCAI, 2011.

[7] G. Fei, A. Mukherjee, B. Liu, M. Hsu, M. Castellanos, and R. Ghosh. Exploiting burstiness in reviews for review spammer detection. In ICWSM, 2013.

[8] A. j. Minnich, N. Chavoshi, A. Mueen, S. Luan, and M. Faloutsos. True view: Harnessing the power of multiple review sites. In ACM WWW, 2015.

[9] B. Viswanath, M. Ahmad Bashir, M. Crovella, S. Guah, K. P. Gummadi, B. Krishnamurthy, and A. Mislove. Towards detecting anomalous user behavior in online social networks. In USENIX, 2014.

January-February 2020
ISSN: 0193-4120 Page No. 10766 - 10768

[10] H. Li, Z. Chen, B. Liu, X. Wei, and J. Shao. Spotting fake reviews via collective PU learning. In ICDM, 2014.

[11] L. Akoglu, R. Chandy, and C. Faloutsos. Opinion fraud detection in online reviews by network effects. In ICWSM, 2013. [12] R. Shebuti and L. Akoglu. Collective opinion spam detection: bridging review networks and metadata. In ACM KDD, 2015.

[12] A. Mukherjee, A. Kumar, B. Liu, J. Wang, M. Hsu, M. Castellanos, and R. Ghosh. Spotting opinion spammers using behavioral footprints. In ACM KDD, 2013.

[13] L. Akoglu, R. Chandy, and C. Faloutsos. Opinion fraud detection in online reviews by network effects. In ICWSM, 2013.

[14] R. Shebuti and L. Akoglu. Collective opinion spam detection: bridging review networks and metadata. In ACM KDD, 2015.

[15] S. Feng, R. Banerjee and Y. Choi. Syntactic stylometry for deception detection. Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Short Papers; ACL, 2012.