Xi'an Shiyou Daxue Xuebao (Ziran Kexue Ban)/
Journal of Xi'an Shiyou University, Natural Sciences Edition
ISSN: 1673-064X
E-Publication: Online Open Access
Vol: 66 Issue 08 | 2023
DOI 10.5281/zenodo.8289267

DETECTION OF SPAM REVIEWS IN WEBSITES USING NET SPAM ALGORITHM

SOWMYA D
Assistant Professor, Computer Science, School of Allied Health Sciences, Mahatma Gandhi Medical College & RI, Sri Balaji Vidyapeeth, (Deemed to be University), Pondicherry, India.

LAVANYA BASKARAN
PG Student, Department of Community Medicine, Indira Gandhi Medical College & RI, Pondicherry University, Pondicherry, India.

UMA AN *
Professor, Medical Genetics, Principal, School of Allied Health Sciences, Mahatma Gandhi Medical College & RI, Sri Balaji Vidyapeeth, (Deemed to be University), Pondicherry, India.
* Corresponding Author Email: umaan@mgmcri.ac.in

Abstract

Now-a-days, people in general choose to purchase an online item depending on the reviews and feedback given in social media. The probability of leaving a scrutiny gives a glorious chance for person who writing a spam reviews regarding the product for individual reasons. Categorizing these kinds of spammers and the spam content create an interesting issue of analysis. Though a generous range of studies are done recently towards this, till date the methodologies used still hardly find the spam reviews. Here, we propose an exclusive framework called Net-Spam that exploit spam options for creating review datasets as heterogeneous data networks to map spam detection procedure into further classifications. Maltreatment of spam options helps us to get the best output pertaining to various metrics experimented on real-time review datasets from Amazon and Yelp websites. The results show that Net-Spam outperforms the current ways from among the three classes of features namely detection of review, detection of user and detection of Spam Groupers.

Index Terms: Social Media, Spammers, Review, Framework, Net-Spam, Heterogeneous data Networks

1. INTRODUCTION

Website portals play basic part in the Generation of information which is avital source for producers and user to advertise to prefer products and services reciprocally. From the previous few years, it is noticed that people are considering the reviews, be it positive or negative [2]. In terms of business, reviews became an important aspect as positive reviews carry benefits whereas negative reviews can cause economic loss [4]. Anybody with any identity can give reviews this provides a convenience for spammers to give fake reviews that deceive the user opinion. By sharing, these negative reviews are duplicated over the net [1]. Reviews that are written for money to change user's attention to buy the product and t also considered to be spam [6]. The generic idea of planned framework is to be given a review dataset as a HIN (Heterogeneous Information Network) and to map spam detection into a HIN classification. Specifically, we have a tendency to model review dataset as a HIN in which reviews are connected through different nodes. Weights are calculated and from these weights have to calculate the ultimate labels of reviews using

Xi'an Shiyou Daxue Xuebao (Ziran Kexue Ban)/
Journal of Xi'an Shiyou University, Natural Sciences Edition
ISSN: 1673-064X
E-Publication: Online Open Access
Vol: 66 Issue 08 | 2023
DOI 10.5281/zenodo.8289267

supervised and unsupervised approaches [7]. Here we used two sample datasets of reviews. Based on investigation, to build two views for options, the classified features as review-detection have more weights and allow higher performance on noticing spam reviews in semi-supervisions build no noticeable variation on the performance of approach [3]. To resolve that feature weights are usually added or removed for labeling and therefore time complexity is extent for a specific level of accuracy. And also use the features with more weights to get high accuracy with less time complexity [5].

2. THEORETICAL ANALYSIS

By using original dataset going to detect whether the review is original or spam. For the better understanding of technique, first present an examination of some of the concepts in heterogeneous information networks.

I. Definitions

Heterogeneous Information Network (HIN):A heterogeneous information network is a graph which is represented as 'H' = (N, E), where every node in the graph and edge belongs to one particular node type and link type reciprocally, 'N' represents nodes and 'E' represents edge i.e., link between two nodes. If two edges are in the same type means, the types of starting node and ending node of those edges are parallel.

Network Schema: A meta-path along with object type mapping and their edge mapping relationship is known as network schema. It generally describes about the number of node types and where the possible edge remain (simply a meta-structure). It is mathematically represented as 'T' = (A, R), where 'A' is object type and relation 'R'.

Meta-path: The arrangement of relationships in network schema is known as a meta-path. It is generally represented in the form M1(R1)M2(R2)...Mn(R (n−1)).It prescribes a complex relation between two nodes and also a meta-path can be characterized by a sequence of node types when there is no uncertainty, i.e., P = M1M2...Mn. There are no edges between two nodes of the same type. The meta-path enhances the approach of edge types to path types and illustrates the different relations among node types through indirect links.

Classification problem: In the HIN, the types of the nodes to be classified and have a few labeled nodes and unlabeled nodes. Classification should be complete to predict the unlabeled nodes. The nodes are classified into different classes C1, C2 …Ck, where 'K' is the number of classes.

II. Feature Types

Here data is the written review. The most of the review is written for increasing the rating value of their product. Based on the user reviews, to identify whether it is fake or honest review .The Metapath is defined through their shared features of the reviews among two nodes. Based on features of users and reviews fall into the different categories as follows:

Xi'an Shiyou Daxue Xuebao (Ziran Kexue Ban)/
Journal of Xi'an Shiyou University, Natural Sciences Edition
ISSN: 1673-064X
E-Publication: Online Open Access
Vol: 66 Issue 08 | 2023
DOI 10.5281/zenodo.8289267

Behavior User based feature: This feature depends on the review of single particular user. To conclude all the written reviews, we need to calculate according to the each individual user review. There are two features in user behavior based on which we can know, even if a review is spam or Original.

1. **Burstiness**: It is calculated based on the time of activity of user and time taken to write review and past review time. So that based on the burstiness value we can know the review is spam or real. Because these kind of spam reviews are written fast.
2. **Negative ratio**: Basically, the opponent gives the ratings as low. Spammers write the reviews to smear the business. So the reviews completely negative with zero rating are spam.

Behavior Review Based Features

This feature depends on the Meta data (data about data). This category has two features.

1. **Early time frame (ETF):** The most spam reviews are on the top so that user visit that review first.
2. **Threshold rating deviation (DEV):** To advertise their business the spammers rate high. So that the mean and variance are high based, on which we can detect the spam messages.

Linguistic User based feature

The assessment and feeling of each particular user are extracted. The spammers commonly write the reviews in the same pattern. There are two features used in this category. They are Average Content Similarity (ACS) and Maximum Content Similarity (MCS). The spammers don't waste their time in writing original review. They write the same reviews. The values calculated for the related reviews lie between 0 and 1.

Linguistic Review based features

In this feature the opinion and feelings of all reviews are examined. In this division spam reviews are identified based on two attributes. They are the ratio of First Personal Pronouns (FPP) and the Exclamation Ratio of Sentences (ERS). Spammers use second personal pronoun that first personal pronoun and also they use! To inspire users. So that the most of the reviews with '!' are acclaimed as spam.

3. METHODOLOGY

Prior/Antecedent Knowledge: Initially compute antecedent knowledge that means review r being spam which denoted as yr. The proposed works are used in both semi-supervised and unsupervised learning. In the semi-supervised method, if review r is labeled as spam yr = 1 in the pre-labeled reviews, else yr = 0. Due to the amount of supervision if the label of their view is unknown, consideryu = 0 i.e., to assume it as a non-spam review. In the unsupervised method, antecedent knowledge is comprehended

Xi'an Shiyou Daxue Xuebao (Ziran Kexue Ban)/
Journal of Xi'an Shiyou University, Natural Sciences Edition
ISSN: 1673-064X
E-Publication: Online Open Access
Vol: 66 Issue 08 | 2023
DOI 10.5281/zenodo.8289267

by using yr= 1/$Lfxl$rLl=1 where fxlr is the probability of review r being spam according to feature l and L denotes number of all the features used.

Network Schema definition: After Antecedent knowledge to define network schema based on a spam factor which determines the features engaged in spam detection. These Schemas are mostly depend upon the basic definitions of metapaths and shows various connections of network components.

Meta path definition and Formation: A Metapath is the package of relations in the network schema. For the formation of Metapath, to define various levels of spam certainty by create one link of network review. The number of Metapath and reviews would be connected to each other through these features by using a higher value increases. Since, enough spam and non-spam reviews for each step with slight numbers of reviews connected to each other for every step, so that the spam probabilities of reviews take uniform distribution but with lower value of enough reviews. Moreover, accuracy for lower levels decreases because of the bipolar problem and it decades for higher values, because they take uniform distribution.

Algorithm: NETSPAMDETECTION ()

Inputs: The given inputs are review dataset, lis tof spam feature, pre-labeled reviews (each review dataset is labeled as spam or real).

Output: importance of features (W) and spamcity Probability (Pr)

r,v-reviews.

#n- number of reviews.

L-number of features.

#$m^{pl}u$,-level of spam inevitability

#priori-knowledge

1.**if** semi-supervised mode

if r €prelabeled-reviews

yr = $labeled$(r)

else

yr = 0

 # determining network schema

2. schema= determine spam_feature_list

#metapathFormation

3. for$pl \in schema$

do: for u,v∈$review\_dataset$

Xi'an Shiyou Daxue Xuebao (Ziran Kexue Ban)/
Journal of Xi'an Shiyou University, Natural Sciences Edition
ISSN: 1673-064X
E-Publication: Online Open Access
Vol: 66 Issue 08 | 2023
DOI 10.5281/zenodo.8289267

do: $m_u{}^{pl}$=|s×$fxlu$ | / s

$m_v{}^{p}$= |s×$fxlu$ | / s

if$m_u{}^{pl}$= $m_v{}^{pl}$

do: $m_{u\ ,v}{}^{pl}$= $m_u{}^{pl}$

else

do: $m_{u\ ,v}{}^{pl}$= 0

#classification: calculation of weight

4. for$pl \in schema$

do:W_{pl} = $\sum_r^n = 1 \sum_s^n = 1 mp^{pl}{}_{r,s} Xy_r$ X y_s

---

$\sum_r^n = 1 \sum_s^n = 1 mp^{pl}{}_{r,s}$

#classification: labeling

5.foru,v∈$datasetreview$

do:$Pr_{u\ ,v}$ = 1 − $\coprod_{pl=1}^{L}$ 1 - $mp^{pl}{}_{u,v} XW_{pi}$

Pr_u = $avg$of $_{,1}$, $Pr_{u\ ,2}$, … … . , $Pr_{u\ ,n}$

6. return (W,Pr)

Classification: The following two steps are used in classification part of NetSpam namely Counting of weight and Labeling.

1. **Counting of weight:** Counting of weight calculates the weight of each and every step of metapath. Based on the similarity of weight with other nodes in the network it is simulated that the classification of nodes is completed with linked nodes of higher probability may have the same labels. The relation in a HIN includes the straight link and the path of the network can be measured by using the perception of metapath. Accordingly, to use the metapath networks defined in the previous step, usually used to calculate the heterogeneous relations between nodes. Further over, the process of the pathwill be able to calculate the weight of separate relation path that will be used to estimate the label of each unlabeled review in the next step.
2. **Labeling:** To build the HIN, used to connect the number of links between a review and other reviews increase the probability of having a label that is related to consider that a node along with other nodes showing their similarities. Moreover, if a review has many links with non-spam reviews, it means features of reviews are shared with other reviews having low spamicity increasing its probability so it may be considered a non-spam review.

Xi'an Shiyou Daxue Xuebao (Ziran Kexue Ban)/
Journal of Xi'an Shiyou University, Natural Sciences Edition
ISSN: 1673-064X
E-Publication: Online Open Access
Vol: 66 Issue 08 | 2023
DOI 10.5281/zenodo.8289267

4. EXPERIMENTAL EVALUATION

In Experimental Evaluation part, discuss about the acquired results based on the dataset and also evaluating the results i.e., whether the proposed approach recognize the spam reviews with high accuracy or not. It is determined by the metrics.

1. **Data Set:** Datasets are collected from Yelp. Yelp is one of the sites where there are reviews for the particular restaurants, hotels, etc. It also advised which the best are. Here examined around408, 600 reviews written by the customer for restaurants and hotels in the city New York. In this dataset few of the reviews are labeled as spam or real. The label is given concede to the yelp algorithm. It is accomplished by yelp recommender, it is not sure that those labels are superlative but they are trustable. The reviews in the data set contain reaction and comments on the quality of the item. The other aspects are the customer id, restaurant name, user ratings, date and time of when the review is written and when the user visited. This main dataset is divided into three other dataset (Table 1) as follows:
    i. Dataset based on reviews: 30% of the data are randomly collected from the main dataset
    ii. Dataset based on Item: 30 % of the data is collected from the dataset with the same item.
    iii. Dataset based on User: Minimal set of reviews of the same user are collected.

Likewise, we have taken another review dataset from the amazon website.

Table 1: Sample Review Datasets

Dataset	Reviews (Spam %)	User	Business (Restaurants & Hotels)
Main	608,600(11%)	300,270	6132
Review-based	93980(11%)	58,212	3,827
User-based	94670(35%)	60,342	4.673
Item-based	110,730(20%)	160293	4,623
Amazon	9,000	8325	356

2. **Results:** To detect the spam reviews by using Net spam algorithm. The outcome of NetSpam algorithm is compared with existing method of other two approaches: random approach and Speagle Plus. The accuracy of each method is compared. And also correlated by the feature weight which was discussed in theoretical analysis part. This framework is analyzed by unsupervised mode and finally the time complexities are also compared.

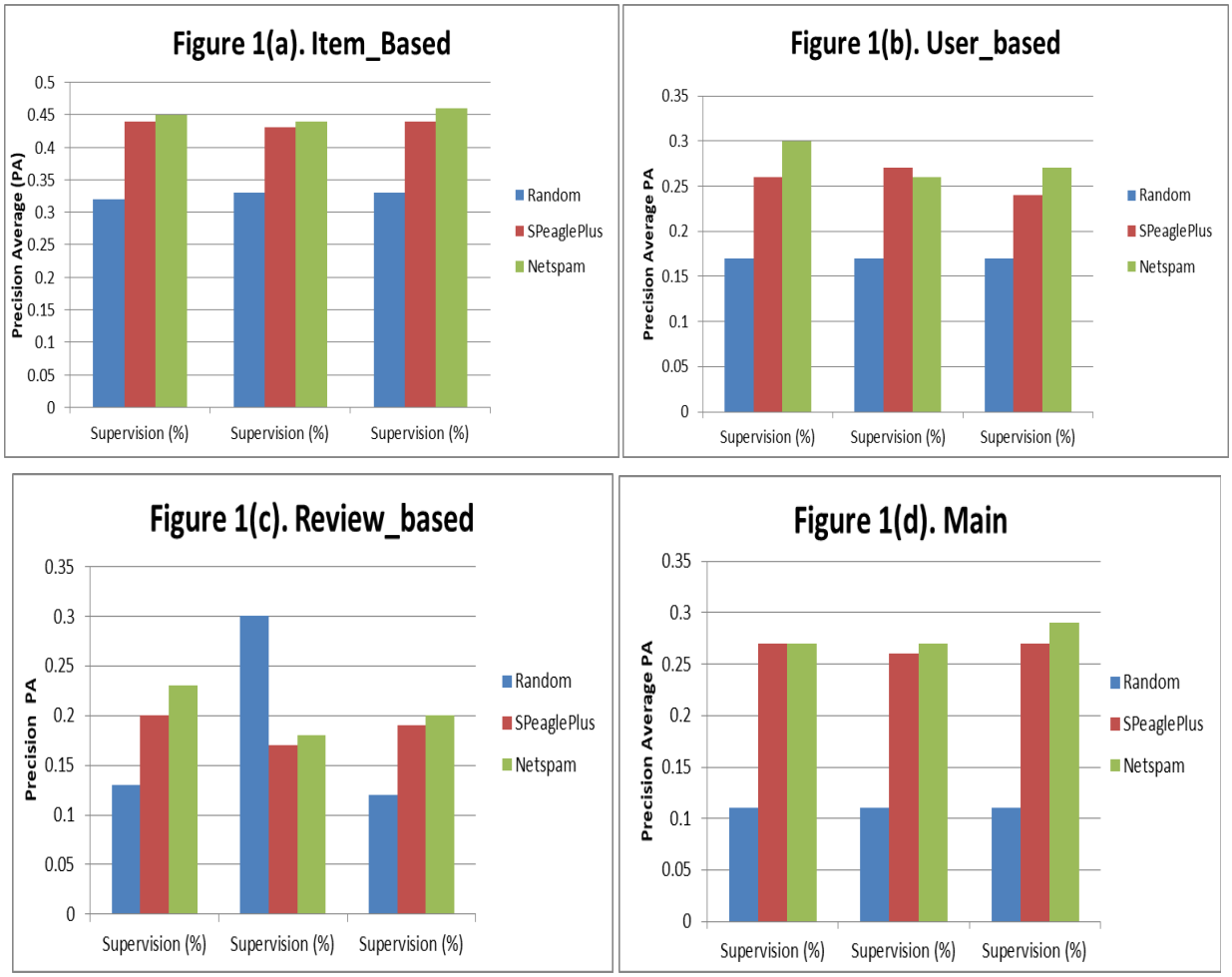
Accuracy: To compare those three approaches, using the Precision Average (PA) and Area under the Curve (AUC). AUC measures the efficiency based on the True Positive Ratio (TPR) against False Positive Ratio (FPR) (Figures.1 a, b, c & d). True positive ratio is real reviews based on positive reviews. Figures 2a, b, c & d represents the AUG values for the different dataset and for different approaches. Although Fig 1 a, b, c & d represents

Xi'an Shiyou Daxue Xuebao (Ziran Kexue Ban)/
Journal of Xi'an Shiyou University, Natural Sciences Edition
ISSN: 1673-064X
E-Publication: Online Open Access
Vol: 66 Issue 08 | 2023
DOI 10.5281/zenodo.8289267

the PA for different datasets in calculating PA, need to sort the spam reviews in the top of the list. The higher index should be spam review and then PA is calculated as follows:

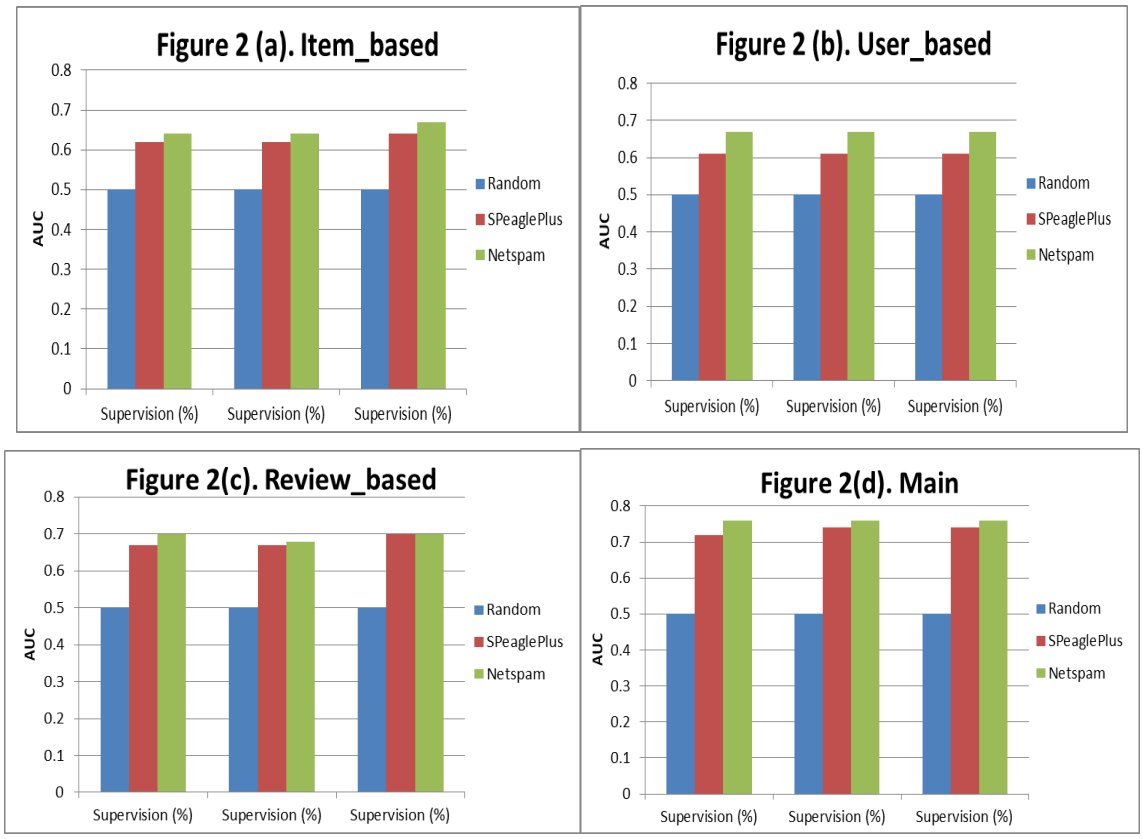
$$PA = \sum_{i}^{n} = 1 \frac{i}{I(i)}$$ where I is index and I is list of dataset.

From the Fig.1 and 2, we observe that NetSpam gives the highest efficiency in detecting the spam review. There is no effect of supervision on the NetSpam and SPealePlus. The PA value based on the spam percentage in the dataset as AUC values do not change.



Figures 1a, 1b, 1c, & 1d: Precision Average (PA) for Random, SPeaglePlus & approaches of Net Spam in various datasets and supervisions (1.5%, 2% and 5.5%)

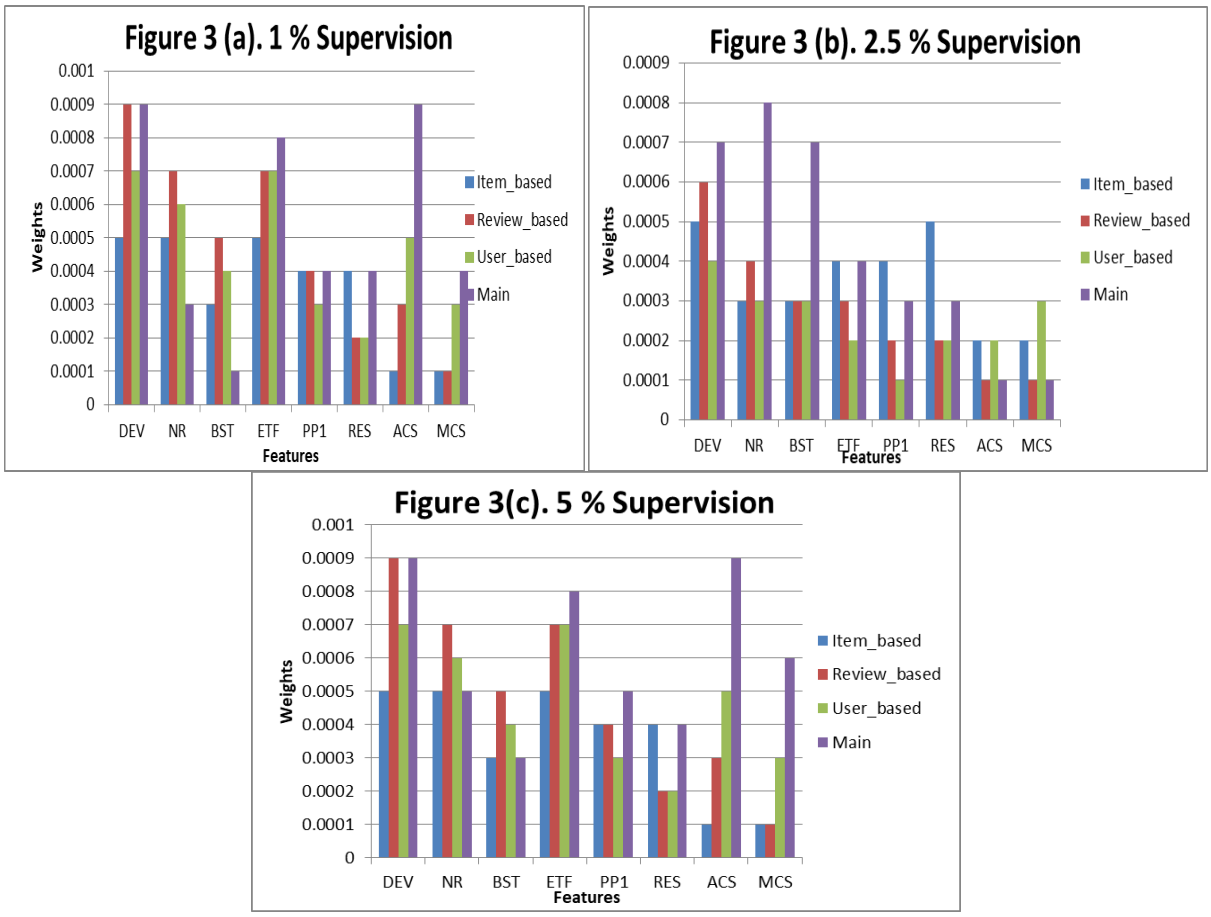
Xi'an Shiyou Daxue Xuebao (Ziran Kexue Ban)/
Journal of Xi'an Shiyou University, Natural Sciences Edition
ISSN: 1673-064X
E-Publication: Online Open Access
Vol: 66 Issue 08 | 2023
DOI 10.5281/zenodo.8289267



Figures 2a, 2b, 2c, & 2d: AUC for Random, SPeaglePlus and Net Spam in various datasets and supervisions (1.5%, 2% and 5.5%)

Feature weight analysis: This analysis conception with the comparison of features of the dataset. So that we can detect spam reviews based on feature with high accuracy. From the figure 3a, b & c, it was observed that the result of the main dataset is ranked first because it contains all the features for every supervision.

Xi'an Shiyou Daxue Xuebao (Ziran Kexue Ban)/
Journal of Xi'an Shiyou University, Natural Sciences Edition
ISSN: 1673-064X
E-Publication: Online Open Access
Vol: 66 Issue 08 | 2023
DOI 10.5281/zenodo.8289267



Figures 3a, 3b & 3c: Features weights for Net Spam framework on various datasets using different supervisions (1%, 2.5% and 5%)

Unsupervised method: In unsupervised method, used to compute basic labels and these labels are used to calculate the feature weight and finally compute review labels. To observe that there is a valuable correlation between the Main dataset in which for Net Spam it is equal to 0.789 (p-value=0.0307) and for SPeaglePlus reach 0.910 (p-value=0.00222).

Time complexity: Time Complexity is time taken to recognize the spam reviews in the offline mode is O (E2m). 'E' is the total number of edges and m is the number of features although in online mode it takes less time approximately O (Em) because in online mode there is no need for repetition in every feature like offline mode.

5. CONCLUSION

This learning suggests a novel based spam detection framework called NetSpam Algorithm based on the concepts of metapath and graph-based method of labeling reviews depending on rank-based labeling. The enforcement of the framework is evaluated by using real datasets of Yelp websites and Amazon websites. By perception

ISSN: 1673-064X
E-Publication: Online Open Access
Vol: 66 Issue 08 | 2023
DOI 10.5281/zenodo.8289267

shows that, using the concept of Meta path calculated weights are very effective in identifying spam reviews and leads to a better performance. Further, Net Spam can able to calculate importance of each feature's and yields better performance in the process of features and performs better than previous works with only a small number of features even without a trained set. Furthermore, after defining four main categories for features observations shows that the behavioral based review category performs better than other categories. The conclusion decides that using different supervisions that are same to the semi-supervised method will not affect the determining features that are most weighted in various datasets. In future proposed approach of Metapath can be applied to other problems. For example, to find the spammer communities in reviews and also community finding based on metapath concept is used to find the features based on reviews of group spammer with highest similarity of challenges. Using the review features is an interesting proposed work that is related to spam reviews and finding spammers. Also, when single network receives scrutiny from various disciplines are used to detect information diffusion and content sharing in multilayer networks is still in research progress.

References

1) Satish Tukaram Pokharkar, Ajit JaysingraoShete, Vishal Dyandeo Ghogare. Survey in Online social media Skelton by network based spam. International Research Journal of Engineering and Technology (IRJET). 2017; 04 (11): 1517- 23.

2) Ch. Xu and J. Zhang. Combating product review spam campaigns via multiple heterogeneous pair wise features. SIAM International Conference on Data Mining (ICDM).2014; 09(1):2422-2431.

3) H. Li, Z. Chen, B. Liu, X. Wei, and J. Shao. Spotting fake reviews via collective PU learning. IEEE International Conference on Data Mining.2014; 8(4):468-474.

4) Vyas Krishna Maheshchandra, Ankit P. Vaishnav.A Survey on Review Spam Detection techniques.I nternational Journal of Engineering Research & Technology (IJERT). 2015:04(4): 368 – 371.

5) H. Xue, F. Li, H. Seo, and R. Pluretti. Trust-Aware Review Spam Detection. IEEE Trustcom/ISPA. 2015; 4(3): 265 – 271.

6) M. Salehi, R. Sharma, M. Marzolla, M. Magnani, P. Siyari, and D. Montesi. Spreading processes in multilayer networks. IEEE Transactions on Network Science and Engineering.2015; 2(2):65–80.

7) K. Amar, M. Kameshwara Rao, Ch. Chaitanya, Ravi Kumar Tenali. A Network-Based Spam Detection Framework for Reviews in Online social media. International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET).2018; 7(2): 1622 – 1627.

8) B. Viswanath, M. Ahmad Bashir, M. Crovella, S. Guah, K. P. Gummadi, B. Krishnamurthy. Anomaly Detection in Online Social Networks: Using Datamining Techniques and Fuzzy Logic. Conference in Queensland University of Technology (CQUT).2014; 4(3): 2341-2349.

9) F. Li, M. Huang, Y. Yang, and X. Zhu. Learning to identify review spam. International Joint Conference on Artificial Intelligence (IJCAI). 2011; 2(4):3041-3048.

10) R. Shebuti and L. Akoglu. Collective opinion spam detection: bridging review network sand metadata. Association for Computing Machinery in Knowledge Data Mining (ACMKDM).
2015; 5 (2):1356-1361.