Journal of
Data Acquisition and Processing

ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.7745025

SPAM REVIEW IDENTIFICATION IN E-COMMERCE PLATFORMS TO HELP E-COMMERCE USERS TO MAKE CORRECT PRODUCT DECISIONS

Nidhi Bhandari
Department of Applied Mathematics, IET, DAVV, Khandwa Road , Indore.
Email: nidhi.bhandari1@gmail.com

Rachna Navalakhe
Department of Applied Mathematics and Computational Science, SGSITS, 23, Park Road, Indore, Email: sgsits.rachna@gmail.com

G.L Prajapati
Department of Computer Engineering, IET, DAVV, Khandwa Road, Indore, India
Email: glprajapati1@gmail.com

Abstract— In e-commerce websites, users are utilizing product reviews to make decisions about a product purchase. But the fake and spam reviews of products are influencing buyers' decisions. In this context, recently different approaches are developed for classifying spam reviews, but most of the authors are only considering only the review text which provides partial information to classify the reviews as spam or legitimate. In this paper, we are proposing a new weighted classification technique to deal with this problem. The proposed technique utilizes the reviewer's attributes and other people's opinions about the review for providing a score for the review. On the other hand, the review text is classified using the text classification approach to decide the initial review class. Further, a weight is calculated using the calculated score and initial class label. Using this weight we have decided on the final class label about the product class i.e. spam or legitimate. The experiments are carried out using the Amazon product review dataset and the performance of the model has been measured. The performance of the prepared model is measured and compared in terms of Precision, Recall, and F1-score. The comparison shows the proposed method provides more accurate results as compared to only text classification-based techniques.

Keywords—Text mining, sentiment analysis, Spam Classification, weighted classifier, ecommerce, decision making process.

I. INTRODUCTION

Text is one of the most used techniques to describe the details of various things in this world. Text is used for providing information, sharing opinions, expressing emotions, and others [1]. Therefore it is used in a number of places in real-world applications. In this context, text processing or automated text handling is a technique to identify the hidden information in a given text block [2]. However, text processing or text mining is a classical domain of research in machine learning. Using text processing a number of applications such as search engines, Information Retrieval (IR), summarization, historical facts digitization, etc are developed [3].

In this presented work, we are demonstrating an application of automated text analysis in order to classify spam text in e-commerce platforms.

In e-commerce platforms, most of the users before purchasing a product read reviews about the products. The product reviews are describing the pros and cons of the product listed in an e-commerce platform [4]. But sometimes e-commerce product vendors are delegating some commercial reviews to attract more buyers to their products for making profits. Therefore these reviews are mostly biased toward the product quality and their actual properties [5]. Such kind of fake reviews are highly influential and can misguide the buyer's decisions. Moreover, such kind of false information not only affecting the buyer's trust it also damages the e-commerce platform's credibility [6]. However, there are some very essential contributions in order to deal with this issue available in literature but most of them are either not much accurate or working on partial information. Therefore classification accuracy is the key issue in spam detection in e-commerce reviews.

In this paper, we are presenting a novel approach to identifying spam reviews by incorporating the review text and the reviewer's profile information to accurately classify the spam reviews. This section provides an overview of the proposed work, the next section provides the details of the proposed weighted classification technique, and then the implemented technique is evaluated and compared with the technique developed based on only text-based spam review classification. Using this comparison we are trying to find the impact of the proposed technique over the traditional methodology of e-commerce spam review classification techniques. Finally, the conclusion is provided and future research directions are discussed.

II. proposed work

The review analysis on an e-commerce platform is most of the time conducted to know the opinion or sentiment of the product users as feedback for the product manufacturer or product vendor. Using such kind of opinion mining the product manufacturers are improving the quality of products to serve better to their consumers [7]. In this perspective, the product manufacturers are trying to identify the sentimental orientation of the product user in terms of positive, negative, or neutral. On the other hand, the same review is manually analyzed by a buyer to make the decision about the product to be purchased or not. But a false or biased review will affect the buyer's decision and results in a bad quality product being delivered to the consumer [8]. That not only affects the consumer's trust but also damage the credibility of the e-commerce platform. Therefore, it is essential to analyze a product review by the e-commerce authorities before publishing it to the platform to maintain a healthy e-commerce platform [9]. Therefore the proposed work is motivated to design a new approach to classify spam reviews in an e-commerce platform. Figure 1 demonstrates the overview of the proposed concept for classifying spam reviews.

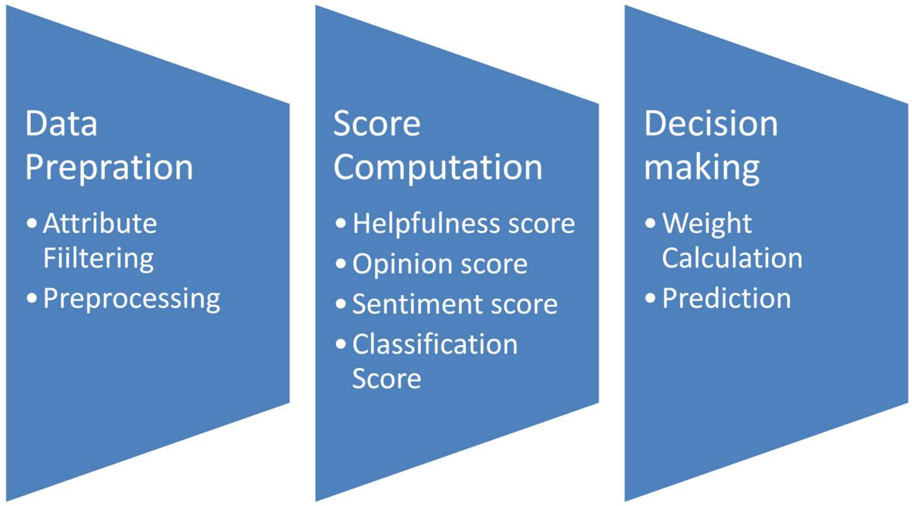


Figure 1 Demonstrate the process of the proposed weight based spam review classification

Dataset: In this context, we have explored different datasets and found an appropriate among them from Kaggle [10]. This dataset is a part of Amazon product reviews and has a number of different categories such as home and kitchen, electronics, toys, games, etc. Among them, we have selected toys and games categories for our proposed model design. Here, the given dataset is available in form of JSON file format. The JSON is a semi-structured XML file format, which consists of product review information according to a specific format. In the first step, we downloaded and parsed the file to prepare two vectors which consist of nine attributes as listed in below table 1.

Table 1 Amazon product review dataset attributes

S. No.	Attribute Name	Type of data
1	Id	Unique ID of review
2	Reviewer Name	Reviewer name who posted the product review
3	votes-down/up	A set of count [down and up]
4	Review Text	Review text post
5	Rating	Rating given by the reviewer
6	Summary	Summary of the review
7	Review Time	Time stamp
8	Category	Product category
9	Class	Spam or legitimate

Data pre-processing: After the data parsing, we need to perform data pre-processing by eliminating unwanted or non-essential data. Therefore first we have eliminated the non-essential attributes namely ID, Review Time, and category. Here, ID is a unique value associated with each review to identify the review individually thus it is an over-fitted attribute, similarly, review time is also a unique value, and therefore, it is necessary to eliminate it for better classification results. Additionally, the category is similar for all the data instances, therefore, it is also not playing any role in the classification task thus we have decided to remove these three attributes. In addition, the reviewer name is also a unique value for all the instances thus we have decided to reduce this attribute also. Thus the final set of attributes is demonstrated in table 2.

Table 2 final set of attributes after data pre-processing

S. No.	Attribute Name	Type of data
1	votes-down/up	A set of count [down and up]
2	Review Text	Review text post
3	Rating	Rating given by the reviewer
4	Summary	Summary of the review
5	Class	Spam or legitimate

Score computation: After finalizing the attributes we utilize the individual attributes for computing four different kinds of scores. The details of the required scores are described below:

Helpfulness score

In review posts, there is a provision to provide an up-vote or down-vote a review post by the other person who is reading the reviews. The up or down vote depends on the criteria and how helpful the review is for making the decision. In this dataset, we have an attribute "votes-down/up". Using this attribute we calculated the helpfulness score using the following formula:

$$H_s = \frac{U}{T} - \frac{D}{T} \dots\dots\dots\dots\dots (1)$$

Where, H_s is helpfulness score, U is the total up votes, D is total down votes, and T is total votes given to the review

$$T = U + D \dots\dots\dots\dots. (2)$$

Opinion score

In different review systems in e-commerce platforms, a rating is also associated. This rating is provided by the reviewer to describe their sentiments towards the product. In most of the platforms, the rating is given between 1-5. In this work the rating of a review is utilized as an opinion score and calculated between 0-1 using the following formula:

$$O_p = \frac{R - M_r}{M_r - M_n} \dots\dots\dots (3)$$

Where, O_p is the opinion score, R is the rating assigned to review, M_r is the maximum rating score and M_n is the minimum value of a rating.

Sentiment score

Sentiment indicates the orientation of the review text. Mostly the sentiments are calculated to classify the text into their sentiment classes such as negative, positive, or neutral. But in this experimental dataset, we don't have any predefined sentiment classes to learn and then estimate the sentiment class for each review text. Therefore, in this presented work we have utilized a popular sentiment score calculation library namely Valence Aware Dictionary and sEntiment Reasoner (VADER) [11]. In order to calculate the sentiment score the library provides a function which is known as "polarity_scores". The polarity score is results in four different scores for the given sentence or paragraph. The type of score and the value range is given in table 3. In this work, we utilize the compound sentiment score for representing the sentiment of the review summary. This compound sentiment score is denoted as C.

Table 3 Sentiment score

S. No.	Sentiment type	Score range
1	Negative	0 - 1
2	Positive	0 - 1
3	Neutral	0 – 1
4	Compound	-1 - +1

Classification score

The next attribute is the review text and the class label. The review text is used to express the quality of a product in the e-commerce platform. Additionally, we have predefined classes to learn and identify the class of a review text. Therefore the review text is used with the text classification technique to identify the class of a given review. In order to classify the review text the required model is demonstrated in figure 2.

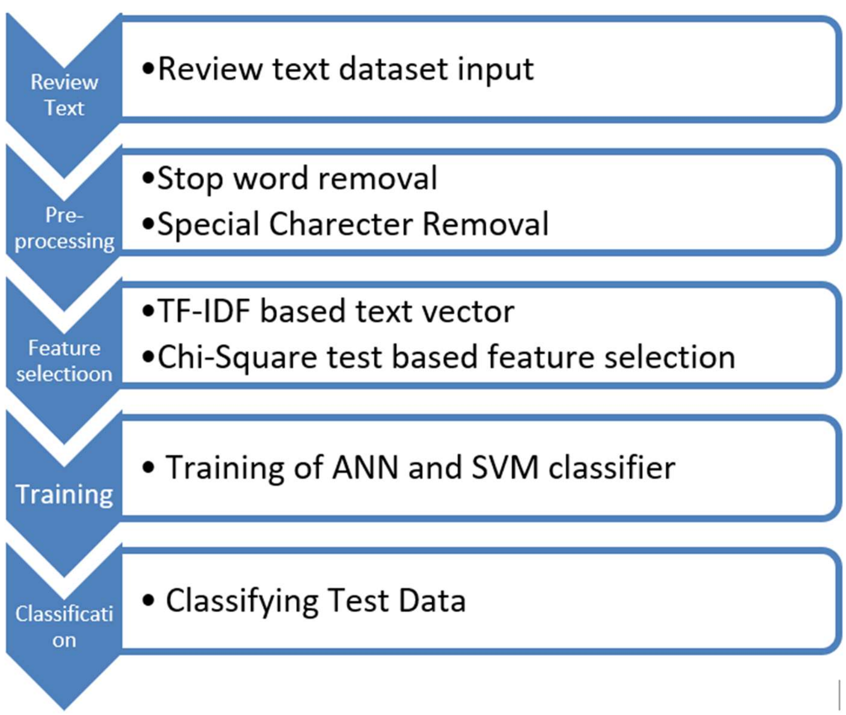
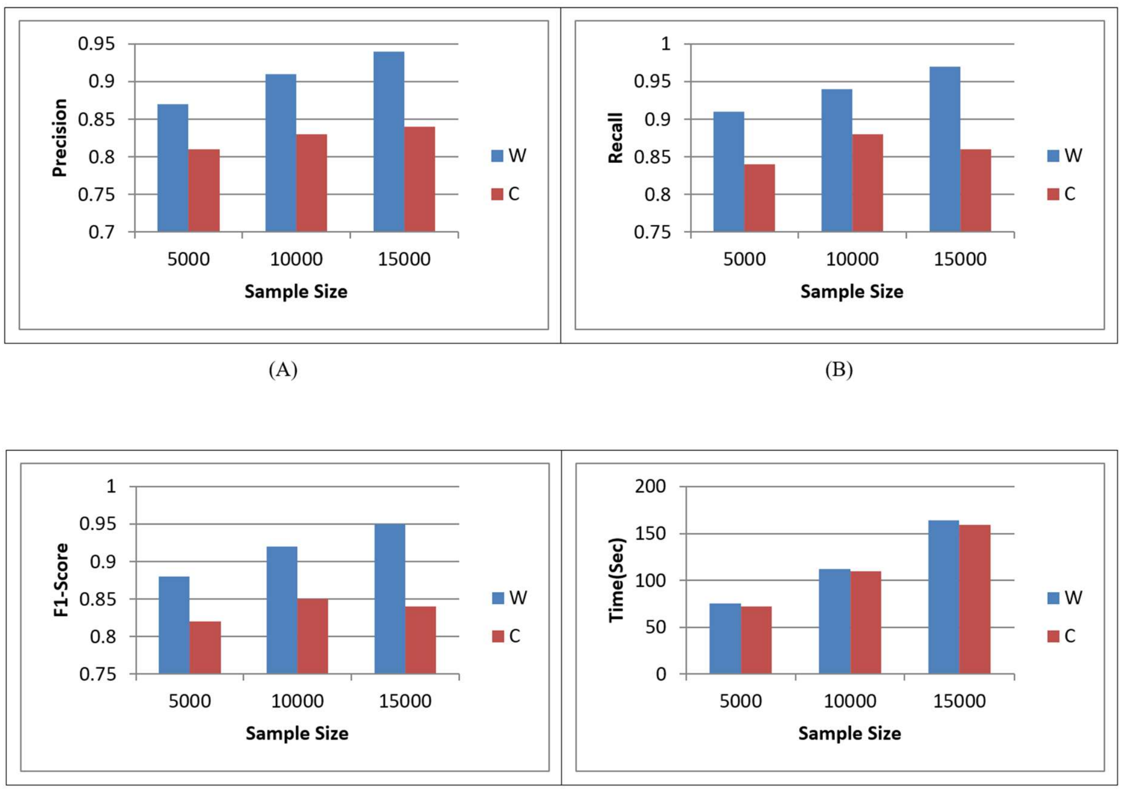


Figure 2 Review Text Classification

In this phase, we have used a part of the dataset containing only the review text and class labels. Then, the data preprocessing technique is used. The preprocessing involves the use of Punctuation Removal, special character removal, and Stop-word Removal. After preprocessing the data we extract the features from the text. Before features extraction, we considered here only two class labels thus we have mapped all the class labels where 1 as spam review and normal text as class 0. Now, we compute Term Frequency and Inverse Document Frequency (TF-IDF). In addition, for reducing the large feature set we utilize the Chi-Square test between the features and the class labels [12]. Using this approach we have selected only 5000 features. Finally, the dataset is split into two parts with a ratio of 75%-25% for training and testing. Both splits are used with the classifiers for performing training and validation. In order to perform training and testing the two classifiers utilized a support vector machine, and an ANN algorithm. The predicted class label is denoted here as L.



(A) (B)

(C) (D)

Figure 3 shows the comparative experimental performance of the proposed weighted classification technique and traditional spam review classification technique in terms of (A) precision (B) recall (C) f-score and (D) training time

Calculating final class label

In order to approximate the final class label for a given instance of review we utilize a weighted technique. The following formula is used for calculating the weight of instance.

$$W = H_s * w_1 + O_p * w_2 + C * w_3 + L * w_4 \dots \dots . . (4)$$

Where the w_1,w_2,w_3 and w_4are the weight factors which is adjusted according to algorithm designer with the constraint of w_1+w_2+w_3+w_4=1. Here we have distributed the same weight to all w_1=w_2=w_3=w_4=0.25.

Now in order to decide the final class label of the data we utilize the following function:

$$f(W,L) = \begin{cases} if\ W \geq 0.66\ and\ L == 0\ then\ legitimate \\ if\ W < 0.66\ and\ L == 1\ then\ legitimate \\ \qquad\qquad otherwise\ \ spam \end{cases} \dots\dots..(5)$$

Finally, the performance of the system has been measured and compared with the SVM based method [13], which is based only on the review text classification. The next section provides a discussion of the performed experiments and obtained results.

experiments and results

In this section, we provide the experimental analysis and obtained comparative results. Therefore, the proposed weighted technique is compared with the review text-based classification technique. The traditional technique utilizes the support vector machine classifier. There are four parameters are evaluated for demonstrating the performance of the proposed technique i.e. precision, recall, f-score, and training time.

The precision can be described as the fraction of relevant instances among the retrieved instances. Precision is defined as follows:

$$precision = \frac{TP}{TP + FP} \dots\dots\dots.(6)$$

Where, TP is true positive and FP is false positive.

Figure 3(A) shows the comparative performance of the proposed weighted spam review classifier and traditional SVM-based review classifier in terms of precision. The precision is also known as the accuracy of a classification system. In this diagram, the X axis shows the experiments with two different sample sizes, and the Y axis shows the precision of the classifier. According to the obtained results, the proposed weighted classification approach provides more accurate results as compared to only content-based review classification. In addition, we also find that when the sample size increases the classification accuracy is reduced but the classification system based on the weighted technique will improve the classification accuracy with increasing size of samples.

Table 4 Observed comparative performance of proposed and traditional spam review classification system

Table 4 Observed comparative performance of proposed and traditional spam review classification system

S. No.	Sample size	Precision		Recall		F1-score		Training time	
		W	C	W	C	W	C	W	C
1	5000	0.87	0.81	0.91	0.84	0.88	0.82	75	72
2	10000	0.91	0.83	0.94	0.88	0.92	0.85	112	110
3	15000	0.94	0.84	0.97	0.86	0.95	0.84	164	159

W = Weighted Classification, C = Content based classification

W = Weighted Classification, C = Content based classification

The next evaluation parameter is Recall, which is also known as sensitivity or true positive rate and is defined as follows:

$$recall = \frac{TP}{TP + FN} \dots\dots\dots\dots.(7)$$

Where TP is true positive and FN is false negative.

Figure 3(B) shows the recall of the proposed and traditional review classification system. In this diagram, the X axis shows the size of the data samples used for the experiment, and the Y axis shows the recall obtained using the proposed weighted classification technique and SVM-based classification technique. According to the obtained results, the proposed technique was found superior then the SVM-based classification technique. Additionally, we have found with the increasing amount of samples the proposed technique provides more consistent results as compared to the SVM classifier which reduces their performance with the sample size.

The next parameter is F1-score. The F1-Score takes into account both precision and recall for measuring the f1-score. Therefore the f1-score is also known as the harmonic mean of precision and recall in order to describe the quality of classification. It is defined using the following formula:

$$F - Score = 2 * \frac{precision * recall}{precision + recall} \dots\dots\dots.(8)$$

Figure 3(C) demonstrates the performance of the classification models in terms of recall. In this diagram, the sample size is given in the X-axis and the Y-axis contains the obtained f1-score. According to the obtained results, the proposed model demonstrates a higher f1-score as compared to the only content-based technique of spam review classification.

Finally, we have measured the time consumption for performing the training of the proposed and traditional model. Figure 3(D) shows the time consumption for performing the training of both models. In this context, X-axis shows the sample size and the Y-axis shows the time required to train the model in terms of seconds (sec). According to the obtained experimental results the SVM takes higher training time as compared to the ANN-based approach. Additionally, when we include the time taken for computing the additional factors and weights to classify the review, the time requirement of the proposed weighted classification technique is increased as compared to only content-based review classification. Finally, the performance of the experiments is also reported in terms of the obtained values which are reported in table 4.

Conclusions

The techniques of text processing and analysis are utilized in a number of different ways to provide benefits to us. Among them, text analysis for the decision-making process is one of the essential tasks. In this presented work text analysis is performed in order to deal with the issue of e-commerce spam review classification. It is one of the critical issues which will be harmful to both i.e. consumers as well as the e-commerce platforms. In this context, a new review analysis methodology has been introduced based on the weight computation of the review. The weight of a review has been calculated based on four different parameters (1) helpfulness score, which indicates how useful a review is to understand the quality or service offered by the e-commerce (2) opinion score, which indicates how other e-commerce users are reacting on the given review in terms of votes (3) sentiment score, based on the review summary we have calculated a sentiment score to indicate the orientation of the given review (4) classification score, in order to calculate this score we performed a classification of review in terms of spam

or legitimate. Finally, these calculated scores are combined using a weight calculation technique to decide whether a given review is spam or legitimate.
The implementation of the proposed technique is performed and their performance has been measured in terms of precision, recall, f-score, and training time. Additionally, the proposed weighted classification technique is compared with a traditional approach of spam review classification based on only review text, extracted text features, and an SVM-based classifier. The comparative results indicate that the proposed technique is providing more accurate results as compared to the traditional spam classification technique. Therefore, we can say the spam review classification not only depends on the review content but it also depends on the opinion of other users and the reviewer's profile. In addition, the classification of spam reviews from legitimate reviews is difficult without considering additional parameters such as rating, opinion, and votes.

References

- K. A. Lindquist, J. K. MacCormack, H. Shablack, “The role of language in emotion: predictions from psychological constructionism”, Front Psychol. 2015; 6: 444.
- H. Sarker, “Machine Learning: Algorithms, Real-World Applications and Research Directions”, SN Computer Science volume 2, Article number: 160 (2021)
- S. A. Alwidian, H. A. B. Salameh, A. N. Alslaity, “Text data mining: a proposed framework and future perspectives”, Int. J. Business Information Systems, Vol. 18, No. 2, 2015
- L. Rong, Z. Weibai, H. Debo, “Sentiment Analysis of Ecommerce Product Review Data Based on Deep Learning”, 4th Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC), 2021 IEEE
- Y. K. Dwivedi, E. Ismagilova, D. L. Hughes, J. Carlson, R. Filieri, J. Jacobson, V. Jain, H. Karjaluoto, H. Kefi, A. S. Krishen, V. Kumark, M. M. Rahman, R. Ramank, P. A. Rauschnabel, J. Rowley, J. Salo, G. A. Tran, Y. Wang, “Setting the future of digital and social media marketing research: Perspectives and research propositions”, International Journal of Information Management Volume 59, August 2021, 102168
- T. Chen, P. Samaranayake, X. Y. Cen, M. Qi, Y. C. Lan, “The Impact of Online Reviews on Consumers’ Purchasing Decisions: Evidence From an Eye-Tracking Study”, Front Psychol. 2022; 13: 865702.
- N. B. V. Le, J. H. Huh, “Applying Sentiment Product Reviews and Visualization for BI Systems in Vietnamese E-Commerce Website: Focusing on Vietnamese Context”, Electronics 2021, 10, 2481.
- P. Nellutla, M. Hudnurkar, S. S. Ambekar, A. D. Lidbe, “Online Product Reviews and Their Impact on Third Party Sellers Using Natural Language Processing”, International Journal of Business Intelligence Research, Volume 12, Issue 1, January-June 2021
- N. Chawla, B. Kumar, “E-Commerce and Consumer Protection in India: The Emerging Trend”, Journal of Business Ethics volume 180, pages581–604 (2022)
- https://www.kaggle.com/datasets/naveedhn/amazon-product-review-spam-and-non-spam
- S. Elbagir, J. Yang, “Twitter Sentiment Analysis Using Natural Language Toolkit and VADER Sentiment”, Proceedings of the International MultiConference of Engineers and Computer Scientists 2019, IMECS 2019, March 13-15, 2019, Hong Kong

- H. B. Tang, L. Zhou, X. Chengjie, Q. Zhu, "A Method of Text Dimension Reduction Based on CHI and TF-IDF", 4th International Conference on Mechatronics, Materials, Chemistry and Computer Engineering (ICMMCCE 2015)
- C. Lai, M. C. Tsai, "An Empirical Performance Comparison of Machine Learning Methods for Spam E-mail Categorization", Proceedings of the Fourth International Conference on Hybrid Intelligent Systems (HIS'04), 0-7695-2291-2/04