

A MACHINE LEARNING APPROACH FOR ANALYZING SPAM PRODUCT REVIEW TO HELP CONSUMERS

Nitin Upadhyay and Dr. Deepika Pathak

Department of Computer Application, Dr. A. P. J. Abdul Kalam University, Indore (M.P.) –
452010

Corresponding Author Email: nitin_upadhyay86@yahoo.com

Abstract:

In e-commerce websites a wide variety of products are available. Additionally, in large e-commerce websites every day a millions of products are introduced. Among them not all the products are very good in quality and services. On the other hand, the clients are mainly depends on product review and rating given on product review and rating section. Therefore the product purchasing or buying decisions are significantly depends on the product rating and review. It is a highly influential factor and can mislead the buyer's decision. In this paper, we proposed a spam review identification model for helping the users by providing spam free and actual reviews of the product. The proposed model will use a text feature extraction technique using TF-IDF and chi-square test. Then, an Artificial Neural Network is trained for identifying the spam product reviews. In order, to conduct the experiments Amazon product review dataset has been used. Additionally a comparative performance study has also been carried out with SVM based classifier to justify the proposed model. According to the obtained performance we have found the proposed model is far superior than the SVM based spam detection model. Keywords: e-commerce, product review, cyber security, fraud, text classification, spam review.

INTRODUCTION

The ecommerce is a new generation technology for offering product and services online. It provides an online showcase of products and services with details. User of ecommerce or end client visits the ecommerce platform and can select appropriate service or product. In order to make buying decision from these platforms the product review and rating is providing a significant role. Most of the online buyers are reading the product reviews and evaluate the rating of the product before purchases. Therefore, the false or spam reviews are influencing the behavior of buyers. In this context, we need an automatic system which classify and/or identify the false reviews of product. In this presented work we are exploring the techniques of spam review classification using machine learning algorithms.

This paper includes a review of the existing approaches to classify spam comments. Using this review we are intended to identify the algorithms, datasets, and results of the current spam review classification systems. Next, we propose a new machine learning model based on text classification to identify fake or spam product reviews using the Amazon product review dataset. Further, the performance of the proposed spam filtering technique is measured in terms of accuracy and time required for performing classification. Finally, the conclusion is provided and future research directions are provided.

PROPOSED WORK

The main aim is to study and design an enhanced spam review detection system for providing correct information about e-commerce product. In this context, the following objectives are established:

1. Exploring and investigating the effective techniques for e-commerce spam detection technique: In this phase a review of existing spam filtering technique has been carried out, additionally the used algorithms and datasets are studied.
2. Developing and enhanced spam detection model for ecommerce product review: Using the literature a technique has been developed for accurately classify the e-commerce product spam and fake reviews.
3. Evaluation of the proposed technique experimentally and perform comparative study: The proposed technique is evaluated and compared against the existing technique for justifying the proposed work.

This section provides the key objective of the proposed work and the next section discuss the review of the existing techniques available for e-commerce spam review classification.

Related Study

The reviews of customers play an essential role in online shopping. People often refer to reviews or comments of previous customers to decide to buy a product. Some people create untruths and illegitimate reviews about the fake quality of products. These reviews are called spam reviews, which confuse consumers and negatively affect online shopping. C. V. Dinh et al [1] propose a dataset called ViSpamReviews, which has an annotation procedure for detecting spam reviews. The dataset consists of two tasks: binary classification and multi-class classification task for identifying the type of spam. The PhoBERT obtained the highest results on both, 88.93%, and 72.17%, respectively.

Table 1 Review summary

Ref.	Work	Dataset	Algorithm
[1]	Spam review detection	ViSpamReviews	PhoBERT
[2]	Review Text mining techniques, such as semantic analysis and sentiment analysis	-	-
[3]	multi-label customer reviews classification	dataset of more than 50,000 reviews in three categories	Different machine learning for multi-label classification
[4]	fake review detection model using Text Classification and Machine Learning	TripAdvisor dataset	Support Vector Machine, K-Nearest Neighbor, and logistic regression
[5]	study the behavior of early reviewers	Amazon and Yelp	reviewers' ratings and helpfulness scores
[6]	fake review detection system	hotels, restaurants, Yelp, and Amazon	CNN-LSTM model

[7]	an approach which integrates content and usage information to detect fake product reviews	Amazon review dataset	fine-grained burst pattern detection
[8]	Notice faux opinions	Syntactic dataset	logistic regression
[9]	discuss research works on fake review	-	Comparison
[10]	Review of spam reviews detection techniques	-	deep learning
[11]	detecting and classifying the reviews as fraudulent/deceptive or truthful	Syntactic dataset	Bidirectional LSTM
[12]	distinguish spam from truthful reviews	Yelp.com – Hotel and Restaurant	EUPHORIA
[13]	spam online reviews detection		summarize the existing network-based approaches

User-generated content (UGC) provides a valuable source of data to understand consumers and intelligent businesses. However, research on UGC for e-commerce applications involves interdisciplinary knowledge, and few studies have summarized the research framework and application directions. S. Li et al [2] derive a framework to summarize the research in this field. Second, widely used text mining techniques are introduced, including semantic and sentiment analysis. They analyze the development of semantic analysis in text representation and semantic understanding. Then, the definition, development, and technical classification of sentiment analysis techniques are introduced. Third, discuss directions of text mining for business, ranging from high-quality UGC detection and consumer profiling to product enhancement and marketing. Finally, research gaps in these efforts are emphasized, and suggestions are provided.

The multi-label customer reviews classification aims to identify different thoughts of customers. Due to the impact of COVID-19, customers have become more prone to shopping online. E. Deniz et al [3] analysis of customer reviews in order to obtain more thoughts about the product, rather than engaging in emotion-based analysis. They form a new customer reviews dataset of reviews to perform the analysis. The created dataset contains more than 50,000 reviews in three categories, and the review has multiple labels according to the comments. They applied machine learning for multi-label classification. Finally, compared and analyzed the results using statistical metrics. As a result, they found a Precision of 0.9157, a Recall of 0.8837, an F1 Score of 0.8925, and Hamming Loss of 0.0278.

Covid-19 and the enforcement of lockdowns, social distancing, and other measures are leading to an increase in online shopping. The significance of online shopping and the use of e-commerce has increased competition between companies. Here, online reviews play a role in boosting a business. Review is a factor in customers’ decision-making. H. Tufail et al [4] proposed a fake review detection model using Text Classification and Machine Learning. They used classifiers such as Support Vector Machine, K-Nearest Neighbor, and logistic regression,

using a bigram model based on pronouns, verbs, and sentiments. This methodology outperforms the yelp dataset and the TripAdvisor dataset with 95% and 89.03% accuracy.

T. Bai et al [5] study the behavior of early reviewers through two real-world large e-commerce, i.e., Amazon and Yelp. They divide product lifetime into three stages, namely early, majority and laggards. A user who has posted a review is considered an early reviewer. The quantitative characterize early reviewers based on their rating behaviors, helpfulness scores received from others, and the correlation of reviews with product popularity. They have found (1) an early reviewer to assign a higher average rating score; and (2) an early reviewer to post more helpful reviews. The analysis indicates that early reviewers' ratings and helpfulness scores influence product popularity. By reviewing post as a multiplayer competition game, they propose a margin-based embedding model. Experiments on two e-commerce datasets have shown that the approach outperforms.

S. N. Alsubari et al [6], developing a fake review detection system for E-commerce. The methodology has used four fake review datasets of hotels, restaurants, Yelp, and Amazon. Further, preprocessing such as stopword removal, punctuation removal, and tokenization have been performed with padding. This method uses different sizes of datasets, n-gram features are developed with a word-embedding layer. Convolutional and max-pooling layers are implemented for dimensionality reduction. Based on gate mechanisms, the LSTM layer is combined with the CNN for learning. Finally, a sigmoid activation function receives the input sequences from the previous layer and performs classification. The CNN-LSTM model was evaluated. The results of the in-domain experiment were 77%, 85%, 86%, and 87% in the terms of accuracy for restaurants, hotels, Yelp, and Amazon. In the cross-domain experiment, the model attained 89% accuracy.

I. Dematis et al [7] propose an approach that integrates content and usage information to detect fake product reviews. This model exploits product reviews and reviewers' behavioral traits by specific spam indicators. This method of fine-grained burst pattern detection is employed to examine reviews generated over "suspicious" time intervals. The reviewer's past history is also exploited to determine the reviewer's overall "authorship" reputation. This approach is validated with the Amazon review dataset. Results show that the method successfully detects spam reviews thanks to the nature of the employed techniques and indicators.

H. Jashmine et al [8] propose a technique to notice such faux opinions through the usage of logistic regression by considering verview-centric elements a typical accuracy of Also, this study illustrates an effect of the "verified purchase" in pretend classification. Large components of the population, i.e. clients study products or save reviews before figuring out what to purchase and whether or not. Because writing faux opinions comes with financial gain, online evaluation websites there has been a large make bigger in elaborate opinion spam. An untruthful review is faux evaluates opinion spam. Positive critiques of an object can entice extra clients and amplify sales; poor critiques of an object can end result in decreased demand and decreased sales. Fake assessment detection has attracted massive interest. However, nonetheless do not filter faux critiques publicly.

J. Kaur et al [9] intended to discuss research works, which are conducted in the field of spam detection by various scholars and give a comparative view of the various techniques used in a recent study.

Y. Ren et al [10], some work based on deep learning has been investigated. However, there is still a lack of a survey, which can analyze and summarize the previous techniques. To address this issue first introduces the task of deceptive opinion spam detection. Then, summarize the existing dataset and construction methods. Third, existing methods are analyzed from two aspects: traditional statistical methods and neural network models. Finally, give some future directions for the task.

A. Salunkhe et al [11] aim at detecting and classifying the reviews as fraudulent/deceptive or truthful. It involves the use of various deep learning techniques for classifying the reviews and an overview of the proposed approach involving Attention-based Bidirectional LSTM to tackle issues related to semantic information in reviews and a comparative study over baseline machine learning techniques for review classification.

G. Andresini et al [12] describe **EUPHORIA** — a novel classification approach to distinguish spam reviews. This approach couples multi-view learning with deep learning, to gain accuracy by accounting for the variety of information associated with reviews' content and behavior. Experiments carried out on two real review datasets from Yelp.com – Hotel and Restaurant – show that the use of multi-view learning can improve the performance of a deep learning for review spam detection. The approach achieves AUC-ROC equal to 0.813 and 0.708 in hotels and restaurants.

In the past decade, many people conducted research on detecting spam reviews using the statistical or deep learning method. L. He et al [13] first introduces the task of spam online reviews detection and makes a common definition of spam reviews. Then, we comprehensively conclude the existing method and available datasets. Third, we summarize the existing network-based approaches in dealing with this task and propose some direction for future research.

Methodology

The proposed work is motivated to analyze the e-commerce review post text in order to identify the spam or misleading posts. In this context, a spam review classification model has been implemented; additionally their overview is presented in figure 1.

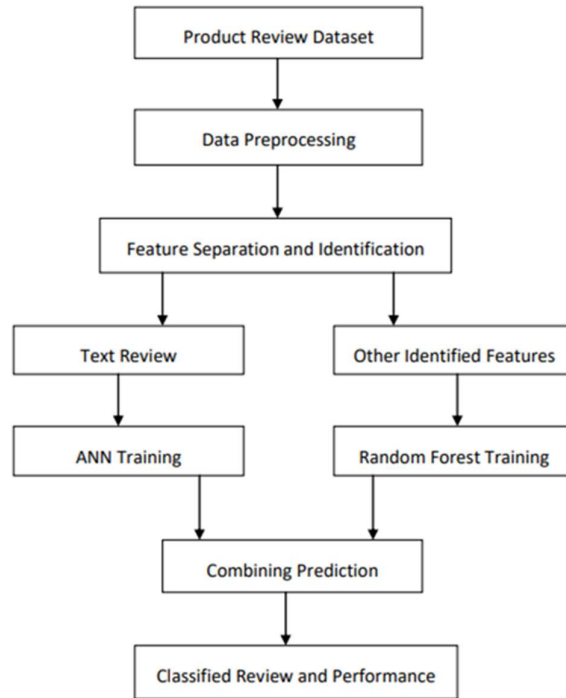


Figure 1 Demonstrate the working flow of the proposed spam review classification model According to the given model, the proposed system accepts the product review dataset. This dataset is taken from the Kaggle repository which is an Amazon product review dataset. The dataset consists of different categories such as electronics, home and kitchen, toys and games, and many others. The dataset is available in the format of JSON files which is an XML format. Among the different categories of products, we have selected the Toys and Games category for performing the experiment. Therefore we first read and parse the JSON files for obtaining the attributes of the review posts. The obtained attributes from the review file are demonstrated in table 2.

Table 2 Dataset attributes

S. No.	Attribute Name	Type of data
1	Id	Unique ID of review
2	Reviewer Name	Review name who reviews the product
3	votes-down/up	The ratio of count [down and up]
4	Review Text	Text review post
5	Rating	Rating given
6	Summary	Summary of the review in form of Text
7	Review Time	Time stamp

8	Category	Product category
9	Class	Spam or legitimate

Now we need to preprocess the data. Preprocessing is a technique by which we remove the non-essential data from the learning information. In this context, we remove some non-essential attributes like ID, Reviewer Name, review time, and category. After the elimination of the additional information, we prepare two subsets of the dataset. This phase of data processing is demonstrated as Feature separation and identification. The two subsets are described as:

Set 1: In this set of information the following attributes are included:

- Votes Down
- Votes Up
- Rating
- Class

Set 2: In this set of information the review text and summary are combined with the class attribute for classification.

Now we have implement two machine learning algorithms for processing both the set of data individually. Therefore, the random forest algorithm has been applied on information set 1. On the other hand for processing of the data we combine the review summary and review text. Next we utilize the Term Frequency and Inverse Document Frequency (TF-IDF) method for extracting the text features from the set 2. This vector of text is further utilized with Artificial Neural Network (ANN) algorithm for performing the training of the algorithm. After training of the algorithms both random forest and ANN the model is become able to predict the spam reviews.

In this context, with the similar ID the test dataset has been prepared. This test dataset is used with the trained algorithms using the trained attributes. Here both the models are predicting the nature of review according to their learning. After getting the prediction from both the algorithms the predictions are combined using the following equation.

$$F_p = 0.4 * R_p + 0.6 * A_p$$

Where, F_p is the final prediction value for the given review R_p is the prediction made by the random forest and A_p is prediction done by the ANN algorithm.

Here the weight of ANN-based prediction is considered higher as compared to the random forest because the ANN provides the prediction based on the content of the review and summary of the review given by the reviewer. On the other hand, the random forest-based prediction considers the attributes of the opinion of others and the reliability of the reviewer. Now the implementation of the proposed concept has been done on the basis of python based technology and their performance analysis was performed. The measured performance of the proposed spam review classification and a traditional SVM-based spam classifier is discussed in the next section.

RESULTS ANALYSIS

The proposed work is aimed to design an e-commerce spam review detection technique. Spam reviews are very influential facts by which the decision of buyers is frequently changed about the product before the product reaches. Therefore accurate product review is essential for a healthy e-commerce system. In this section, we have tested the spam product review detection technique and compared it with a traditional approach of SVM-based spam detection. In this context, we have considered the two key performance parameters namely accuracy and training time. The performance analysis of the system is performed based on the following experimental scenarios:

Providing the performance of random forest classifier with the additional information of review

Providing the performance of ANN algorithm based on the review content

Providing the performance of combined prediction and compared with the SVM based classification approach

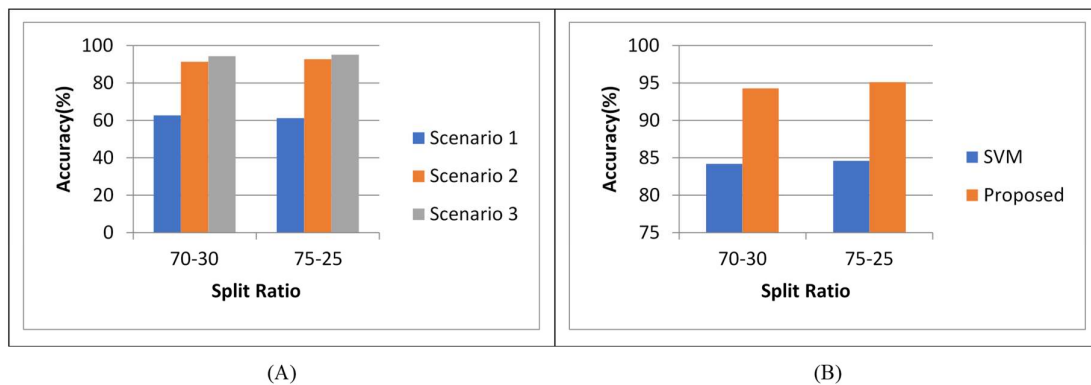


Figure 2 Accuracy of the proposed experimental model in terms of (A) three different scenarios (B) comparing with SVM

Figure 2 Accuracy of the proposed experimental model in terms of (A) three different scenarios (B) comparing with SVM

Here the accuracy describes the correctness of spam identification by the trained classifiers. In our case the accuracy can be defined as a ratio of total correctly identified spam reviews over total reviews provided for classification using the following equation:

$$accuracy = \frac{Correctly\ classified\ spam\ review}{total\ reviews} \times 100$$

Additionally, the training time is also measured, which is an essential parameter of measuring resource consumption. That indicates the amount of time required to train the algorithm using the given amount of data. The time consumption can be measured using following equation:

$$time\ consumed = End\ time - Start\ Time$$

The accuracy of the implemented experimental scenarios are given in figure 2(A) and the comparison with the SVM classifier is given in figure 2(B). According to the obtained accuracy as demonstrated in figure 2(A) we found that the individual accuracy for classifying the additional review parameters provides lower accuracy as compared to content based review

classification. Moreover, the combination of additional review features with the review content enhances the accuracy of the spam identification. On the other hand, the performance reported on figure 2(B) shows the proposed technique is far superior than the traditional SVM based classification approach.

On the other hand the performance of same experimental scenarios in terms of training time is demonstrated in figure 3. The figure 3(A) demonstrates the training time of the implemented scenarios of experiments. According to the obtained performance the simple additional feature based classification requires very fewer amount of time as compared to the content based approach and the combined feature based approach. Here, the combined feature based approach and content based approach consumes similar amount of time. Similarly the performance given in figure 3(B) shows comparative training time of the proposed approach and traditional technique of SVM classifier for spam review classification. Based on the obtained performance the proposed technique found efficient as compared to traditional SVM based spam review classification approach. Thus the proposed methodology is efficient as well as accurate for spam review classification on e-commerce platform.

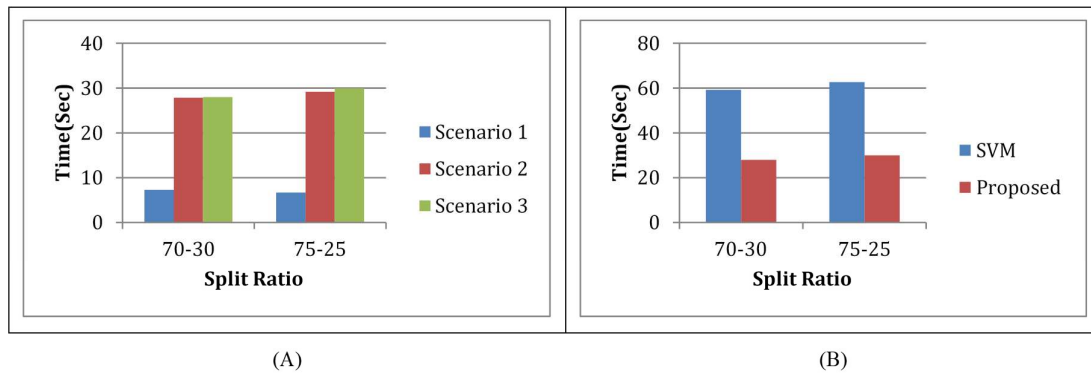


Figure 3 shows training time of the implemented e-commerce spam review classification in terms of (A) individual scenarios (B) comparing with the SVM based classification approach

CONCLUSION & FUTURE WORK

E-commerce is a growing application in recent years. This e-commerce is also including a number of new consumers and product vendors every day. In addition, the introduced product by new vendors is not always effective and good in quality. But due to competition in online business and to increase the revenue of online sales online product vendors are posting spam and fake reviews about the products. However, fake or spam reviews are increasing the revenue of the product but this is not a good sign of a healthy e-commerce application. Additionally, by using these spam reviews the buyers are influenced by the false or bad quality of products. Therefore, spam review identification and elimination is an essential task for maintaining the reputation of the e-commerce platform.

In this context, the proposed work is motivated to design a novel e-commerce spam review classification system. This system not only considers product review text content it also utilizes the reviewer's background and the opinion provided by other product buyers in terms of votes up and down. The system implements two different classification techniques namely random

forest classifier and artificial neural network. The random forest classifier is used to classify the reviewer's information and ANN is used to classify the review content. Finally, both predictions are combined using the weighted technique to produce a final prediction about the review post in terms of spam and legitimate.

Finally, the performance of the implemented system has been measured in terms of accuracy and training time. According to the obtained performance, we found as compared to individual review processing or reviewer's information processing the combined approach is providing more accurate classification results. These results are also verified in comparison with the traditional SVM spam classification approach. According to the comparative performance, we found the proposed technique offers more accurate results as compared to SVM-based e-commerce review classification.

REFERENCES

- C. V. Dinh, S. T. Luu, A. G. T. Nguyen, "Detecting Spam Reviews on Vietnamese E-commerce Websites", arXiv:2207.14636v1 [cs.CL] 27 Jul 2022
- S. Li, F. Liu, Y. Zhang, B. Zhu, H. Zhu, Z. Yu, "Text Mining of User-Generated Content (UGC) for Business Applications in E-Commerce: A Systematic Review", *Mathematics*, 2022, 10, 3554
- E. Deniz, H. Erbay, M. Cosar, "Multi-Label Classification of E-Commerce Customer Reviews via Machine Learning", *Axioms* 2022, 11, 436.
- H. Tufail, M. U. Ashraf, K. Alsubhi, H. M. Aljahdali, "The Effect of Fake Reviews on e-Commerce During and After Covid-19 Pandemic: SKL-Based Fake Reviews Detection", *VOLUME 10*, 2022
- T. Bai, W. X. Zhao, Y. He, J. Y. Nie, J. R. Wen, "Characterizing and Predicting Early Reviewers for Effective Product Marketing on E-Commerce Websites", *IEEE Transactions on Knowledge and Data Engineering*, VOL. , NO. , AUGUST 2018
- S. N. Alsubari, S. N. Deshmukh, M. H. Al-Adhaileh, F. W. Alsaade, T. H. H. Aldhyani, "Development of Integrated Neural Network Model for Identification of Fake Reviews in E-Commerce Using Multidomain Datasets", *Hindawi Applied Bionics and Biomechanics* Volume 2021, Article ID 5522574, 11 pages
- I. Dematis, E. Karapistoli, A. Vakali, "Fake Review Detection via Exploitation of Spam Indicators and Reviewer Behavior Characteristics", *SOFSEM 2018, LNCS 10706*, pp. 581–595, 2018.
- H. Jashmine, K. Sneha, R. Swetha, U. Ruchitha, R. Dhamodharan, "Detection of Online Fake Review in E-Commerce", *International Journal of Innovative Research in Engineering*, Volume 3, Issue 3, May-June 2022, PP: 278-282
- J. Kaur, "A Study on Spam Detection with Sentiment Analysis", *Advances and Applications in Mathematical Sciences*, Volume 21, Issue 5, March 2022, Pages 2695-2705
- Y. Ren, D. Ji, "Learning to Detect Deceptive Opinion Spam: A Survey", *IEEE Access*, *VOLUME 7*, 2019
- A. Salunkhe, "Attention-based Bidirectional LSTM for Deceptive Opinion Spam Classification", arXiv:2112.14789v1 [cs.CL] 29 Dec 2021

G. Andresini, A. Iovine, R. Gasbarro, M. Lomolino, M. de Gemmis, A. Appice, “EUPHORIA: A neural multi-view approach to combine content and behavioral features in review spam detection”, *Journal of Computational Mathematics and Data Science* 3 (2022) 100036

L. He, X. Wang, H. Chen, G. Xu, “Online Spam Review Detection: A Survey of Literature”, *Human-Centric Intelligent Systems* (2022) 2:14–30