GCMK: Detecting Spam Movie Review Based on Graph Convolutional Network Embedding Movie Background Knowledge

Hao Cao[1 [0000−0002−4452−3853]], Hanyue Li[1 [0000−0002−0554−167X]], Yulin He[1 [0000−0001−6842−6970]], Xu Yan[2 [0000−0002−4096−9575]], Fei Yang[1 [0000−0001−9163−7872]], and Haizhou Wang[1,⋆ [0000−0003−1197−5906]]

[1] School of Cyber Science and Engineering, Sichuan University
[2] College of Computer Science, Sichuan University
{caohao, lihanyue, heyulin, yanxu, Yangfei}@stu.scu.edu.cn,
whzh.nc@scu.edu.cn

Abstract. In recent years, the movie industry is booming and many consumers regard movie reviews as an important reference for choosing movies. In the mean time, more and more movie marketing teams glorify their movies and suppress rival movies of the same period by hiring spammers to publish massive misleading reviews. It results in the existence of a large number of spam movie reviews on the online movie review platforms, which greatly misleads consumers and seriously undermines the healthy development of the movie industry. At present, there is little research on spam movie reviews, whose method of detecting spam movie reviews mainly relies on the text and statistical features of reviews, ignoring the significance of movie background knowledge such as movie characters and plots. In this paper, we propose a novel method for detecting spam movie reviews, which uses a graph convolutional neural network embedding movie background knowledge (GCMK). Specifically, we firstly construct a directed heterogeneous knowledge graph by using the movie synopses and the high-quality long comments of movies. Then we use the graph convolutional neural network to obtain the embedded features of the movie background knowledge, use BERT (Bidirectional Encoder Representations from Transformers) model to extract the text features of reviews, and obtain the correlation vectors between these reviews and the corresponding movies by comparing the embedded features of the movie background knowledge and the text features of reviews. Finally, we fuse text features, user statistical features and correlation vectors to construct the detection model. The experimental results demonstrate our proposed GCMK method is more effective than the other state-of-the-art baselines, with an F1-score of 84.94%.

Keywords: Movie reviews · Spam detection · Movie background knowledge · Graph embedding · Features fusion.

⋆ Corresponding author: whzh.nc@scu.edu.cn (H. Wang)

1 Introduction

1.1 Background

With the continuous development of social economy, watching movies has become an important part of people's entertainment. According to the annual report[3] from the Motion Picture Association, the global box office revenue reached $21.3 billion in 2021. At the same time, with the rapid development of Internet technology, more and more people publish and share movie reviews on online movie ticketing and review platforms (such as Maoyan[4], Douban[5], IMDB[6], Rotten Tomatoes[7]). Movie reviews and ratings have become important references for consumers to make choices on purchasing movie tickets [17]. Existing research shows that positive word of mouth of movies (including reviews and ratings) from popular online movie ticketing and review platforms will have a significant impact on movie box office [3, 16]. Therefore, driven by commercial interests, a growing number of spam movie reviews are utilized to influence the movie box office on major online movie ticketing and review platforms [3]. On one hand, this phenomenon seriously misleads consumers who are going to watch movies; on the other hand, it also leads movie marketing team to invest enormous amount of money in false propaganda and other improper manipulations of movie reviews, which greatly damages users' consumption experience and undermines the healthy development of the movie industry. Hence it is extremely necessary and urgent to conduct research on the detection of spam movie reviews.

1.2 Challenges

In recent years, online movie reviews have drawn more and more attention with the popularity of online movie ticketing and review platforms. Compared with the online reviews of e-commerce [4,6,19,21] and restaurant [2,4,11,12], there is very rarely research on spam movie reviews due to its recent emergence. Although spam movie reviews are similar to the spam reviews on the online e-commerce and restaurant platforms, there are some significant differences. (1) Movie reviews often involve specific movie plots, actors, and other movie background knowledge. (2) Movie reviews can be evaluated from more abundant perspectives such as plots, special effects, actors' acting skills and movie types [3, 15]. (3) The life cycle of movies is shorter compared with other products [16], so spammers will focus on the first couple of weeks after the movie released to manipulate the reputation of the movie, thereby affecting the movie box office. Due to the differences mentioned above, the existing research methods for spam reviews cannot be directly applied to the detection of spam movie reviews.

[3] https://www.motionpictures.org/wp-content/uploads/2022/03/MPA-2021-THEME-Report-FINAL.pdf
[4] https://www.maoyan.com/
[5] https://www.douban.com/
[6] https://www.imdb.com/
[7] https://www.rottentomatoes.com/

At present, the main challenges of the research work on the detection of spam movie reviews are as follows: (1) As an emerging research field, there is no available benchmark dataset for researchers (Note that the only research work [3] for the detection of spam movie review did not release their dataset). (2) The existing detection model [3] for spam movie reviews ignores the fact that the content of the movie review should be closely related to the movie background knowledge such as the plot, characters, and actors. As a result, the performance of the existing model for spam movie reviews needs to be further improved. (3) Because of the differences between spam movie reviews and other spam reviews, the existing detection models for spam reviews cannot be directly transferred to the detection research of spam movie reviews, and thus new detection models against spam movie reviews need to be built.

1.3 Contributions

In order to address above problems, this paper builds and publishes a benchmark dataset on GitHub website[8] (Maoyan Dataset), and proposes a **G**raph **C**onvolution-based **M**ovie background **K**nowledge embedding model (GCMK[9]) to detect spam movie reviews. To be specific, we first collect large-scale data from one of the largest Chinese online movie ticketing and review platforms (i.e. Maoyan) and then construct a large-scale spam movie review dataset named M-Dataset using web crawlers. After that, a directed heterogeneous knowledge graph is constructed using the movie synopses and the high-quality long movie comments. The embedding features of the movie background knowledge are obtained by the graph convolutional neural network. Then, we use BERT (Bidirectional Encoder Representations from Transformers) [1] model to extract the text features of the reviews and obtain the correlation vectors between the reviews and the movies by comparing the movie background knowledge embedding vectors and the text feature vectors. Finally, we fuse text features, user statistical features, and correlation vectors to construct a spam movie review detection model.

The main contributions of this paper include the following aspects:

- To the best of our knowledge, we are the first to release large-scale benchmark dataset for detection of spam movie reviews. Specifically, we crawled a huge amount of movie review data from Maoyan (one of the largest online movie ticketing and review platforms in China), covering 2,352 movies of almost all genres, with 734,130 review records in total. A spam movie review dataset called M-Dataset containing a total of 65,696 reviews was constructed by manual annotation, which has been publicly available[8].
- We propose a novel feature extraction method for movie background knowledge based on graph convolutional neural network. First, we construct a directed heterogeneous graph containing movie synopses and real high-quality

[8] https://github.com/yiyepianzhounc/M-Dataset
[9] https://github.com/yiyepianzhounc/GCMK

long comments. Then, we use a graph convolutional neural network to obtain the embedded features containing movie background knowledge such as movie characters and plots. Moreover, we calculate the correlation vectors between the reviews and the corresponding movies by comparing the graph embedding vectors with the feature vectors of the review text. Finally, they are fused with other features to support spam movie review detection.
- We propose a novel spam movie review detection model GCMK based on deep neural network, which integrates user statistical features, text features and movie background knowledge features. Research results on large-scale experimental dataset demonstrate that our proposed model can effectively combine the above three features, improving accuracy of detecting spam movie reviews and outperforming the state-of-the-art baselines.

The rest of the paper is organized as follows. Section 2 will go through related work and achievements in the field of spam review detection. Our proposed GCMK model is elaborated in Section 3. In Section 4, we present the experimental setup and evaluation results. Finally, conclusions are drawn and future research directions are outlined in Section 5.

2 Related Work

Spam movie reviews are essentially a branch of spam reviews. Early research on spam review detection mainly focused on reviews in e-commerce [4,6,19,21], restaurants [2,4,11,12] and other industries [2,14]. The existing detection methods include supervised, unsupervised and semi-supervised learning methods. Among them, models based on supervised learning have been widely used in the detection of spam reviews. The methods [6, 11, 20] mainly apply machine learning to classify online reviews by manually extracting features. However, manual feature extraction is too time-consuming and laborious, so the existing research work [2, 4, 12, 14, 21] et al. proposed various spam review detection methods based on deep learning models, which overcomes the shortcoming that manual features can not extract the deep semantics of reviews. However, supervised learning needs to label a large amount of original data, which is also time-consuming. Besides, the model accuracy is also susceptible to the quality of labeling. Therefore, in parallel with the development of supervised learning methods, research work [3,19] used the unsupervised or semi-supervised methods to identify spam reviews.

Although a number of early studies have focused on movie reviews, they mainly concern about mining opinions in movie reviews for sentiment classification [9,18], summarization [10,13] and making box office predictions [16,17,24]. Currently, Gao et al. [3] is the only research work to study the detection of spam movie reviews. To be specific, Gao et al. [3] adopted an unsupervised learning method based on generative adversarial neural network, which is user-centric and learns textual and statistical features to distinguish the authenticity of movie reviews. However, they ignore the importance of the correlation between user reviews and movie background knowledge for the judgment of the authenticity of reviews. Moreover, they do not release a publicly available benchmark dataset.

3 Methodology

In this section, we describe our method for spam movie review detection based on graph convolution-based movie background knowledge embedding.

3.1 Dataset Construction

Currently, Maoyan is one of the largest online platforms for movie ticketing and review in China. In our research, we choose Maoyan platform to collect movie reviews to construct our dataset. Based on referring existing annotation methods[10] for spam reviews and our rigorous annotation, we construct a benchmark dataset as shown in Table 1. Among them, M-Raw-Data (Maoyan Review Raw Data) is the original collected review records without data annotation. The M-Dataset (Maoyan Review Dataset) is the benchmark dataset of review records after data cleaning and labeling. We finally obtain 65,696 review annotations from 457 movies, including 20,092 spam review records.

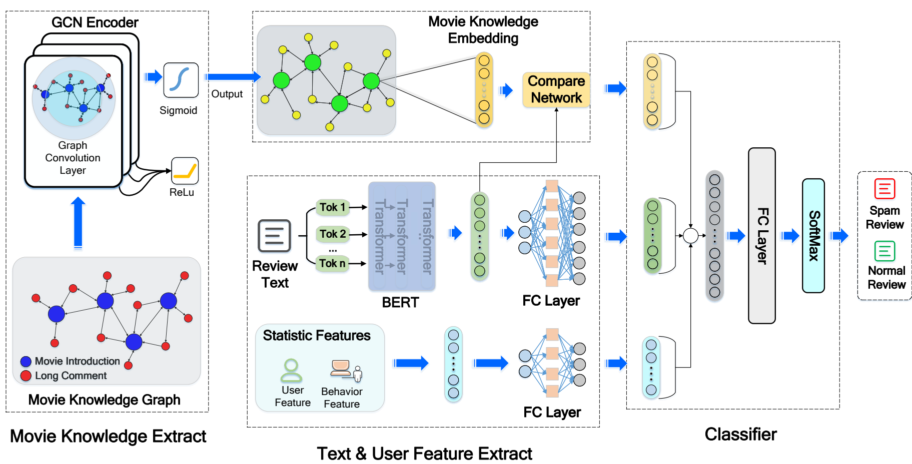

Fig. 1. GCMK Model Structure

3.2 Detection Model

As shown in Fig. 1, we propose a spam movie review detection model that uses graph convolutional neural network with movie background knowledge embedded. Firstly, we extract text features and user statistical features to obtain text

[10] https://consumerist.com/2010/04/14/how-you-spot-fake-online-reviews/

Table 1. Dataset Statistics

Name	#Movie	#Reviews	#Spam	#Non-spam	Genre	Movie Released Year
M-Raw-Data	2,352	734,130	-	-	40	2017-2021
M-Dataset	457	65,696	20,092	45,604		

feature vectors and user statistical feature vectors. Secondly, we use graph convolutional neural network to extract the embedded features of movie background knowledge, and then obtain the correlation vectors between reviews and movies through the comparison network. Finally, the three feature vectors are concatenated and fed into the fully connected layer and the Softmax layer to obtain the classification results.

Movie Background Knowledge Extraction

Building a Movie Background Knowledge Graph We construct a directed heterogeneous graph $\delta = (V,\ E)$ that contains movie synopses and high-quality long comments of movies. The graph contains two types of nodes: movie synopses $S = \{V_1^s, V_2^s, ...,\ V_i^s, ..., V_M^s\}$ and long comments $L = \{V_1^l, V_2^l, ...,\ V_j^l, ..., V_N^l\}$, where M denotes the number of movie synopsis nodes and N denotes the number of long comment nodes. To be specific, V_i^s denotes the ith movie synopsis node, V_j^l denotes the jth long comment node, i.e., $V = S \cup L$. And the set of edges E include bidirectional connections and unidirectional connections. The details of constructing the graph are described as follows.

The set of M movies is denoted as $F = \{f_1, f_2, f_3, ..., f_M\}$, and one movie can belong to multiple themes. Movies of the same theme may have certain similarities in the plots, which leads to some similarities of user reviews. Firstly, we establish a bidirectional connection between the movie synopsis nodes of the same theme movies, so as to allow useful background knowledge to propagate among the same theme movies. Secondly, we establish a bidirectional connection between the long comment node and the movie synopsis node in the same movie, and expand the useful movie background knowledge through multiple high-quality long comments from the reviewers. Finally, in a long comment, the reviewer tends to compare the movie they reviewed with other movies, including the quality of the movies, the plots, the characters, etc. So we establish a connection between the long comment node and the other movie synopses. In order to avoid embedding other unrelated movie information into movie background knowledge through long comment nodes, the edges we build here are unidirectional edges.

Constructing Directed Heterogeneous Graph Convolutions Based on the above-constructed directed heterogeneous movie background knowledge graph $\delta = (V, E)$, we design a graph convolutional neural network based on the directed heterogeneous graph. First, we use the pre-trained Chinese-BERT model to embed the node features. The vector $x_s \in R^D$ denotes the embedded feature of the node V_s. The matrix $X = \{x_1, x_2, x_3, ..., x_{|V|}\} \in R^{|V| \times D}$ contains the feature vectors of all nodes, and each row of X represents the feature vector of one node. We define A is the adjacency matrix and D is the degree matrix. Then, the heterogeneous graph convolution layer updates the $(i+1)$th layer representation of the aggregated features by aggregating the features of their adjacent nodes.

$$A' = D^{-\frac{1}{2}} (A + I) D^{-\frac{1}{2}} \tag{1}$$

$$H^{(i+1)} = \sigma \left(A' H^i W^i \right) \tag{2}$$

where I is the identity matrix of dimension $|V|$. The A' is the adjacency matrix after self-connection and normalization. The W^i is the weight matrix of the ith layer, The H^i is the feature matrix of the ith layer, σ is the nonlinear activation function (such as the $ReLU$ function), and $H^{(i+1)}$ is the next layer after aggregation feature matrix.

Feature Extraction of Review Correlation After obtaining the embedded representation of the movie background knowledge, we compare the review feature vector T_n with the embedded vector x_n of the corresponding movie background knowledge, and get their correlation vector $S_n = f_{cmp}(T_n, x_n)$, where $f_{cmp}()$ is the comparison function. Based on [5], we design the comparison function as $f_{cmp}(x, y) = W [x - y, x \odot y]$, where W is the transformation matrix, x and y are the movie background knowledge embedding vector and the review feature vector, respectively, $\odot$ denotes element-wise multiplication for two matrixes.

Text and User Statistical Feature Extraction

Text Feature Extraction First, we use the review r_n as the input data to the pre-trained Chinese-BERT model to get its vector representation, and then feed the vector representation into the fully connected layer to get the text feature vector T_n.

Table 2. User Statistical Features

Feature Name	Description
User Rating (UR)	It reflects the user's rating towards the movie.
Movie Rating (MR)	It reflects the audience's average rating of the movie.
Consensus of Opinion (CO)	Using $\|UR - MR\|$ to evaluate the consistency of opinions between one reviewer and other audience.
Account Active Degree (ACD)	The level of activity of the account making the review.
Review Time (RT)	The user's review time in a day.
Time Span (TS)	The span between review time and movie release time.

User Statistical Feature Extraction According to the characteristics of movie reviews and the Maoyan platform, we propose six statistical features of users. The details are shown in the Table 2.

3.3 Feature Concatenation

First, we concatenate the text feature representation vector T_n, the user statistical feature vector U_n and the comment correlation vector S_n to get the vector F_n, where $F_n = concat(T_n, U_n, S_n) \in R^{|S_n|+|U_n|+|T_n|}$. Then the vector F_n is fed into the Softmax classification layer, which is simplified as $Z = Softmax(WF_n + b)$ where W is the parameter matrix of the fully connected layer, and b is the bias vector of the fully connected layer.

4 Experiments

In this section, we will evaluate the performance of our proposed GCMK method. First, we describe the experimental setup in our work. Then, we illustrate the superiority of our method in comparison with the baseline methods. Then, the effects of different modules on the model are examined by ablation experiments. Finally, we verify the robustness of our model by randomly adding different proportions of noise to the training set.

4.1 Experiment Settings

We implemented our GCMK approach and baseline model with Pytorch v1.10.1 framework. In our work, all experiments were performed on a workstation equipped with Intel (R) Core (TM) i9-10900 CPU and NVIDIA GeForce RTX 3070 GPU with 64GB of memory.

We use the M-Dataset for training and verification. Specifically, we do not change the number of spam reviews, randomly sample the same number of spam reviews from 45,604 non-spam reviews, and construct a balanced dataset containing 40,184 reviews. We reproduce three models for comparison experiments, where addCGAN [3] is the first state-of-the-art model for spam movie review detection, and CBRNN [21] and HACL [14] are the state-of-the-art models for spam review detection. At the same time, we select a couple of very popular text classification models like TextCNN [22], FastText [8], DPCNN [7], AttBiLSTM [23] as our baseline models. Four evaluation metrics are used to quantify the effect of the models, including Accuracy, Precision, Recall, and F1-score.

4.2 Baseline Model Comparison Experiment

To demonstrate the effectiveness of our proposed GCMK method, we test our model and seven other baselines on our built M-Dataset. The results are shown in Table 3. The results show that our proposed GCMK method outperforms other baseline models on spam movie review detection in terms of Accuracy, Precision, Recall and F1-score. Compared with these state-of-the-art detection methods, GCMK is proved to effectively extract more discriminative features of spam movie reviews (background knowledge of movies and user features) and achieves significant performance improvements in detecting spam movie reviews.

Table 3. Results for Baseline Model Comparison Experiment

Model	Accuracy	Spam			Non-Spam		
		Precision	Recall	F1-score	Precision	Recall	F1-score
TextCNN [22]	74.80	77.77	70.31	73.85	72.29	77.00	79.41
FastText [8]	76.73	80.01	73.41	76.42	73.79	80.55	77.02
DPCNN [7]	77.14	77.51	77.26	77.39	76.75	77.00	76.88
AttBiLSTM [23]	76.77	79.50	74.06	76.68	74.25	79.67	76.86
CBRNN [21]	76.77	77.91	76.71	77.31	75.60	76.84	76.22
HACL [14]	72.83	72.83	71.82	72.45	72.59	73.83	73.21
addCGAN [3]	76.08	76.95	75.96	76.45	75.19	76.21	75.70
GCMK(our model)	**85.01**	**83.22**	**86.73**	**84.94**	**86.85**	**83.38**	**85.08**

4.3 Feature Ablation Experiment

Our proposed GCMK model combines text features, user statistical features and the movie background knowledge. We explore the contribution of each proposed block by ablation experiment. Specifically, the use of ablation features is shown in Table 4, where K denotes movie background knowledge, U denotes user statistical feature. And we use Accuracy, Precision, Recall and F1-score to measure the detection effect of the model. The experimental results are shown in Fig. 2. The results show that our designed movie background knowledge embedding and user statistical feature extraction are better than merely using text features. Meanwhile, each block plays a role in the effectiveness of the GCMK model, and the ablation of both block can weaken the effect of the model. These two blocks give a comprehensive improvement to the GCMK model. Evidently, our proposed blocks can effectively help distinguish spam movie reviews.

Table 4. The Description of Feature Sets

Feature set	Categories of features included
GCMK	Movie background knowledge, User statistical feature, Text feature
GCMK$\backslash K$	User statistical feature, Text feature
GCMK$\backslash U$	Movie background knowledge, Text feature
GCMK$\backslash K\&U$	Text feature

4.4 Robustness Experiment

We randomly label the training set incorrectly according to a certain ratio (5%-45%), and train our model and the baseline models to test the robustness under

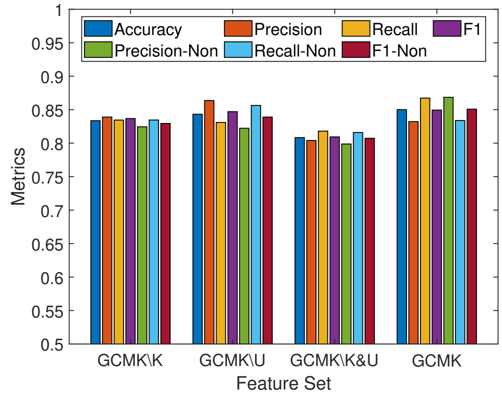

Fig. 2. Results for Ablation Experiment

different noise levels. The experimental results are shown in Fig. 3. The results show that as the noise rate continues to increase, the performances (F1-score) of all models almost decrease to varying degrees, while our model changes little.

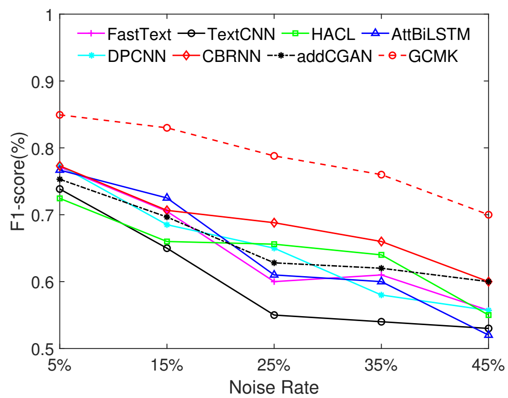

Fig. 3. Results for Robustness Experiment

5 Conclusion

In this paper, we construct the first publicly available Chinese spam movie review dataset (M-Dataset) for spam movie reviews detection, and innovatively propose to use graph convolutional neural network to extract movie background knowledge to help spam movie review detection. A novel GCMK model is developed with movie background knowledge embedded to truly make progress in the problem of spam movie reviews. It utilizes graph convolutional neural network to get the movie background knowledge features based on a directed heterogeneous graph and uses the text features to capture the correlations of reviews and the corresponding movie. Futhermore, we extract six features about spam movie reviews in terms of user behavior and user information. We combine movie background knowledge with text features and user statistical features to improve spam movie reviews detection. The experimental results on our M-Dataset demonstrate our proposed method outperforms the other state-of-art models.

In the future, we will attempt to introduce the correlation comparison algorithm between reviews and the reviewed objects into other review detection fields (e-commerce, restaurant, etc.).

6 Acknowledgments

This work is supported by the National Natural Science Foundation of China (NSFC) under grant nos. 61802271, 61802270, 81602935, and 81773548. In addition, this work is also partially supported by Joint Research Fund of China Ministry of Education and China Mobile Company (No. CM20200409), Sichuan University and Yibin Municipal People's Government University and City Strategic Cooperation Special Fund Project (No. 2020CDYB-29), Science and Technology Plan Transfer Payment Project of Sichuan Province (No. 2021ZYSF007) and The Key Research and Development Program of Science and Technology Department of Sichuan Province (No. 2020YFS0575, No.2021YFG0159, No. 2021KJT0012-2021YFS0067).

References

1. Cui, Y., Che, W., Liu, T., Qin, B., Yang, Z.: Pre-training with whole word masking for chinese bert. IEEE/ACM Transactions on Audio, Speech, and Language Processing **29**, 3504–3514 (2021)
2. Fang, Y., Wang, H., Zhao, L., Yu, F., Wang, C.: Dynamic knowledge graph based fake-review detection. Applied Intelligence **50**(12), 4281–4295 (2020)
3. Gao, Y., Gong, M., Xie, Y., Qin, A.K.: An attention-based unsupervised adversarial model for movie review spam detection. IEEE Transactions on Multimedia **23**, 784–796 (2020)
4. Hajek, P., Barushka, A., Munk, M.: Fake consumer review detection using deep neural networks integrating word embeddings and emotion mining. Neural Computing and Applications **32**(23), 17259–17274 (2020)
5. Hu, L., Yang, T., Zhang, L., Zhong, W., Tang, D., Shi, C., Duan, N., Zhou, M.: Compare to the knowledge: Graph neural fake news detection with external knowledge. In: Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing. pp. 754–763 (2021)
6. Jindal, N., Liu, B.: Opinion spam and analysis. In: Proceedings of the 8th International Conference on Web Search and Data Mining. pp. 219–230 (2008)
7. Johnson, R., Zhang, T.: Deep pyramid convolutional neural networks for text categorization. In: Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics. pp. 562–570 (2017)

8. Joulin, A., Grave, E., Bojanowski, P., Mikolov, T.: Bag of tricks for efficient text classification. arXiv preprint arXiv:1607.01759 pp. 1–5 (2016)
9. Kaur, G., Sharma, A.: Has: Hybrid analysis of sentiments for the perspective of customer review summarization. Journal of Ambient Intelligence and Humanized Computing pp. 1–14 (2022)
10. Khan, A., Gul, M.A., Zareei, M., Biswal, R., Zeb, A., Naeem, M., Saeed, Y., Salim, N.: Movie review summarization using supervised learning and graph-based ranking algorithm. Computational Intelligence and Neuroscience **2020**, 1–14 (2020)
11. Li, J., Ott, M., Cardie, C., Hovy, E.: Towards a general rule for identifying deceptive opinion spam. In: Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics. pp. 1566–1576 (2014)
12. Li, L., Qin, B., Ren, W., Liu, T.: Document representation and feature combination for deceptive spam review detection. Neurocomputing **254**, 33–41 (2017)
13. Liu, C.L., Hsaio, W.H., Lee, C.H., Lu, G.C., Jou, E.: Movie rating and review summarization in mobile environment. IEEE Transactions on Systems, Man, and Cybernetics **42**(3), 397–407 (2011)
14. Liu, Y., Wang, L., Shi, T., Li, J.: Detection of spam reviews through a hierarchical attention architecture with n-gram cnn and bi-lstm. Information Systems **103**, 101865 (2022)
15. Liu, Z., Sun, Q., Yang, Z., Jiang, K., Yan, J.: Fined-grained aspect extraction from online reviews for decision support. In: 2020 IEEE 19th International Conference on Trust, Security and Privacy in Computing and Communications. pp. 1536–1541 (2020)
16. Ma, H., Kim, J.M., Lee, E.: Analyzing dynamic review manipulation and its impact on movie box office revenue. Electronic Commerce Research and Applications **35**, 100840 (2019)
17. Ma, H., Kim, J.M., Lee, E.: A 2020 perspective on "analyzing dynamic review manipulation and its impact on movie box office revenue". Electronic Commerce Research and Applications **41**, 100950 (2020)
18. Manek, A.S., Shenoy, P.D., Mohan, M.C., et al.: Aspect term extraction for sentiment analysis in large movie reviews using gini index feature selection method and svm classifier. World Wide Web **20**(2), 135–154 (2017)
19. Neisari, A., Rueda, L., Saad, S.: Spam review detection using self-organizing maps and convolutional neural networks. Computers & Security **106**, 102274 (2021)
20. Ott, M., Choi, Y., Cardie, C., Hancock, J.T.: Finding deceptive opinion spam by any stretch of the imagination. In: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies. pp. 309–319 (2011)
21. Soubraylu, S., Rajalakshmi, R.: Hybrid convolutional bidirectional recurrent neural network based sentiment analysis on movie reviews. Computational Intelligence **37**(2), 735–757 (2021)
22. Zhang, Y., Wallace, B.: A sensitivity analysis of (and practitioners' guide to) convolutional neural networks for sentence classification. arXiv preprint arXiv:1510.03820 (2015)
23. Zhou, P., Shi, W., Tian, J., Qi, Z., Li, B., Hao, H., Xu, B.: Attention-based bidirectional long short-term memory networks for relation classification. In: Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics. pp. 207–212 (2016)
24. Zhou, Y., Zhang, L., Yi, Z.: Predicting movie box-office revenues using deep neural networks. Neural Computing and Applications **31**(6), 1855–1865 (2019)