Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 4, April : 2023

DETECTION OF FAKE ONLINE REVIEWS USING SUPERVISED AND SEMI SUPERVISED LEARNING

Ms. CH. YAMINI [1], Ms. C. REEMA BLESSY [2], Ms. A. KEERTHI REDDY [3],Ms. N. LAKSHMI SUJITHA [4], MRS. J. KANNAMMA [5]

1. BTECH, VIJAYA INSTITUTE OF TECHNOLOGY FOR WOMENS, ENIKEPADU, VIJAYAWADA, ANDHRA PRADESH, INDIA – 521108 yaminichimata3@gmail.com

2. BTECH, VIJAYA INSTITUTE OF TECHNOLOGY FOR WOMENS, ENIKEPADU, VIJAYAWADA, ANDHRA PRADESH, INDIA – 521108

3. BTECH, VIJAYA INSTITUTE OF TECHNOLOGY FOR WOMENS, ENIKEPADU, VIJAYAWADA, ANDHRA PRADESH, INDIA – 521108

4. BTECH, VIJAYA INSTITUTE OF TECHNOLOGY FOR WOMENS, ENIKEPADU, VIJAYAWADA, ANDHRA PRADESH, INDIA – 521108

5.Asst. Professor COMPUTER SCIENCE AND ENGINEERING, VIJAYA INSTITUTE OF TECHNOLOGY FOR WOMENS,ENIKEPADU, VIJAYAWADA, ANDHRA PRADESH-521108 kannapersis@gmail.com

ABSTRACT

User reviews can have a big impact on how much money a company makes in e-commerce. Before choosing any goods or service, online users rely on reviews. Therefore, the validity of internet evaluations is essential for organisations and has a direct impact on their reputation and portability. Because of this, some companies pay spammers to publish phoney reviews. These fraudulent reviews take advantage of consumer purchasing choices.As a result, during the past twelve years, a lot of research has been done on how to spot false reviews. However, a survey that can evaluate and enumerate the current methods is still lacking. This survey article summarises the existing data sets and their techniques of data collecting in order to address the issue and describes the task of fake review identification. It examines the currently used feature extraction methods. In order to find gaps, it also critically summarises and analyses the available methodologies into two groups: deep learning techniques and conventional statistical machine learning techniques. Additionally, we carry out a benchmark research to assess the effectiveness of various transformers and neural network models that have not previously been used to the identification of fraudulent reviews. The experimental results on two benchmark datasets demonstrate that Robert an outperforms state-of-the-art approaches in a mixed domain for the deception dataset with the maximum accuracy of 91.2%, which may be utilised as a baseline for further research. We conclude by highlighting the research's present limitations as well as potential future possibilities.

1 INTRODUCTION

Customers can share their ratings or thoughts on several websites in current internet age. These reviews are beneficial to businesses and potential customers who may use them to learn more about goods or services before making a decision. It has been noted that there have been a lot more consumer reviews in recent years. Potential customers' decisions are influenced by customer feedback. In other words, consumers decide whether to buy a product after reading reviews on social media or to change their minds. Consumer reviews therefore provide an important service to people.

Positive evaluations result in significant financial gains, while negative reviews frequently have an adverse psychological impact. Customer views are increasingly being used to transform businesses by improving their products, services, and marketing as a result of consumers becoming less responsive to the market. For instance, the maker of an Acer laptop was motivated to create a higher-resolution variant when several customers who bought the laptop left reviews criticising the poor display quality.

Positive evaluations result in significant financial gains, while negative reviews frequently have an adverse psychological impact. Customer views are increasingly being used to transform businesses by improving their products, services, and marketing as a result of consumers becoming less responsive to the market. For instance, the maker of an Acer laptop was motivated to create a higher-resolution variant when several customers who bought the laptop left reviews criticising the

poor display quality. Similar to this, the producer can instruct the paid individuals to publish critical remarks about rival companies' goods in order to promote the business. Reviews that are provided by persons who haven't used the products themselves are regarded as phoney reviews. Therefore, a person who posts false reviews is referred to as a spammer.A group of spammers is the term used when one spammer collaborates with several other spammers to accomplish a particular objective.

The issues of fake review detection have been the subject of numerous investigations. Classifying a review as false or real is the primary task involved in fake review detection. In order to better pinpoint current issues for potential future approaches in this field of study, we have offered a thorough literature review in this survey report.

It offers both deep learning and conventional statistical machine learning methodologies to help researchers interested in fake review identification select the most effective machine learning approach. In this research, relevant publications from Google Scholar, Web of Sciences, and various prestigious conferences are presented to illustrate the difficulties in the field of fake review detection. The papers from 2007 to 2021 that have been identified for summary and analysis are finally.

2. LITERATURE SURVEY AND RELATED WORK

1 Revisiting Semi-Supervised Learning for Online Deceptive Review Detection, J. K. Rout, A. Dalmia, and K.-K. R. Chop, Opinion reviews have an economic influence on businesses' bottom lines as more consumers use online reviews to help them make decisions about services. It should come as no surprise that shady individuals or groups have tried to take advantage of or manipulate online opinion reviews (such as spam reviews) in order to profit or achieve other ends, and that identifying dishonest and phoney opinion reviews is a subject of current study interest. Using a data set of hotel reviews, we demonstrate the usefulness of semi-supervised learning algorithms for spam review detection in this work.

2 Detecting product review spammers using rating behaviors, E. P. Lim, V.-A. Nguyen, N. Jindal, B. Liu, and H. W. Lauw, The purpose of this study is to identify users that create spam reviews or review spammers. In order to identify review spammers, we find a number of distinctive behaviours and model them. We aim to emulate the following behaviours in particular. First, in order to have the most possible impact, spammers may target particular products or product categories. Second, they frequently rate things differently from the other reviews. We suggest scoring techniques and use them on an Amazon review dataset to determine how much spam each reviewer has contributed. Using a web-based spammer evaluation tool created specifically for user evaluation trials, we then choose a selection of highly suspect reviewers for closer examination by our user evaluators. We suggest scoring techniques and use them on an Amazon review dataset to determine how much spam each reviewer has contributed. Using a web-based spammer evaluation tool created specifically for user evaluation trials, we then choose a selection of highly suspect reviewers for closer examination by our user evaluators.

3 Towards a general rule for identifying deceptive opinion spam, J. Li, M. Ott, C. Cardie, and E. Hovy, Online user reviews are increasingly having an impact on consumers' shopping decisions. As a result, there has been an increase in concern about the possibility of posting misleading opinion spam, which consists of made-up evaluations that have been produced with the intent of fooling the reader. Based on a new gold standard dataset that includes data from

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 4, April : 2023

three distinct domains (e.g., hotel, restaurant, doctor), each of which contains three types of reviews, i.e., customer generated truthful reviews, Turker generated deceptive reviews, and employee (domain-expert) generated deceptive reviews, we explore generalised approaches for identifying online deceptive opinion spam in this paper. We anticipate that our method can aid customers when making purchasing decisions and review portal operators, such as Trip Advisor or Yelp, examine any fraudulent behaviour on their sites. Our method seeks to capture the broad difference of language usage between deceptive and accurate reviews.

4 Finding deceptive opinion spam by any stretch of the imaginationM. Ott, Y. Choi, C. Cardie, and J. T. Hancock, More consumers are rating, reviewing, and researching things online. As a result, websites that host user reviews are increasingly being targeted by opinion spam. In contrast to current research, which has mainly concentrated on manually identifying instances of opinion spam, in this work we explore misleading opinion spam, which is made up information that has been purposefully created to appear genuine. We design and compare three methods for identifying dishonest opinion spam by fusing research from psychology and computational linguistics. We then create a classifier that is almost 90% accurate on our benchmark opinion spam dataset. We also provide various theoretical contributions based on feature analysis of our learnt models, elucidating a connection between erroneous beliefs and innovative writing.

5 Detection of review spam: a surveyA. Heydari, M. A. Tavakoli, N. Salim, and Z. Heydari, Online reviews have emerged as the primary source for consumer feedback in recent years. More people and businesses are using these reviews to guide their purchase and business decisions. Unfortunately, fake (spam) reviews have been created by con artists motivated by greed for money or attention. Because of the fraudsters' actions, enterprises that are changing the way they do business and potential customers are misled, and opinion-mining systems are unable to produce reliable results. The current study focuses on categorising and methodically analysing algorithms that identify review spam. The study then goes on to evaluate them in terms of accuracy and outcomes. We discover that studies may be divided into three groups that concentrate on ways to identify spam reviews, specific spammers, and collective spam. Different detection methods favour various detecting scenarios since they each have unique strengths and drawbacks.

3 Implementation Study

1) Data collection: We will upload the AMAZON reviews dataset to the programme using this module.

2) Data Pre-processing: Using this module, we will read all of the reviews, remove stop words, special characters, punctuation, and numerical data from all of the reviews, and then apply Pre-processing to all of the reviews to extract features.

3) Features Extraction: To transform string reviews into numeric vectors, we will use the TF-IDF (term frequency Inverse Document Frequency) technique. Words will be replaced by vectors that contain each word count.

4) Execute the SVM algorithm: We will use the TF-IDF vector to train the SVM algorithm, and then test data will be applied to the trained SVM model to determine the SVM prediction accuracy.

5) Execute the Nave Bayes algorithm: We will use the TF-IDF vector to train the algorithm, and then test data will be applied to the trained model to determine the accuracy of the prediction made using the Nave Bayes algorithm.

6) Execute the Decision Tree method: We will use the Decision Tree method on the TF-IDF vector to train it, and we will use test data on the trained Decision Tree model to determine its prediction accuracy.

4 PROPOSED WORK

Each review in the suggested system first goes through a tokenization procedure. After that, extraneous words are eliminated, and potential feature words are created.

Each potential feature word is compared to the dictionary to determine whether it has an entry. If it does, its frequency is calculated and added to the column in the feature vector that corresponds to the word's numeric map.

The length of the review is calculated and added to the feature vector together with counting frequency.

The feature vector is then updated to include the sentiment score that is included in the data set. Positive sentiment has been given some positive value in the feature vector, while negative sentiment has a zero value.

Advantage of proposed work:

Due to semi-supervised and supervised learning, the system is incredibly quick and efficient.centred on the review-based techniques' substance. Word frequency, emotion polarity, and review length were employed as characteristics.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 4, April : 2023

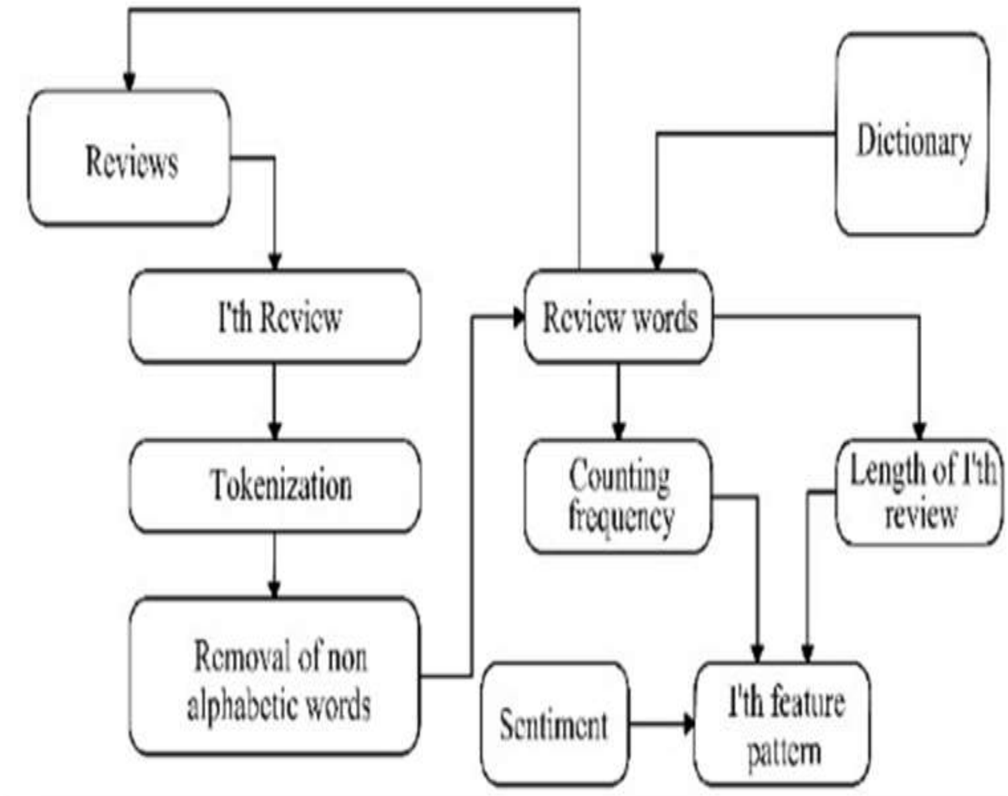

Fig-1: System Architecture.

5 RESULTS AND DISCUSSION SCREENSHOTS

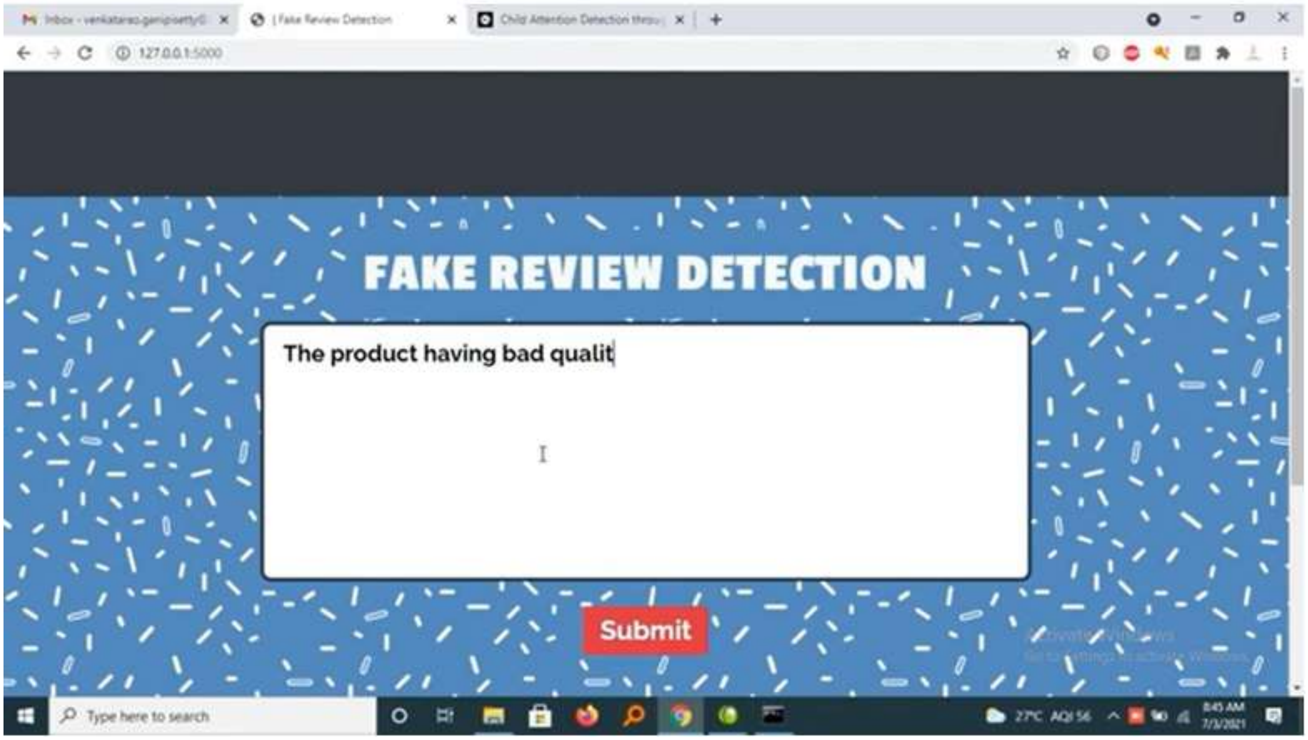

Fig-2: Input-1 Enter the review for detection

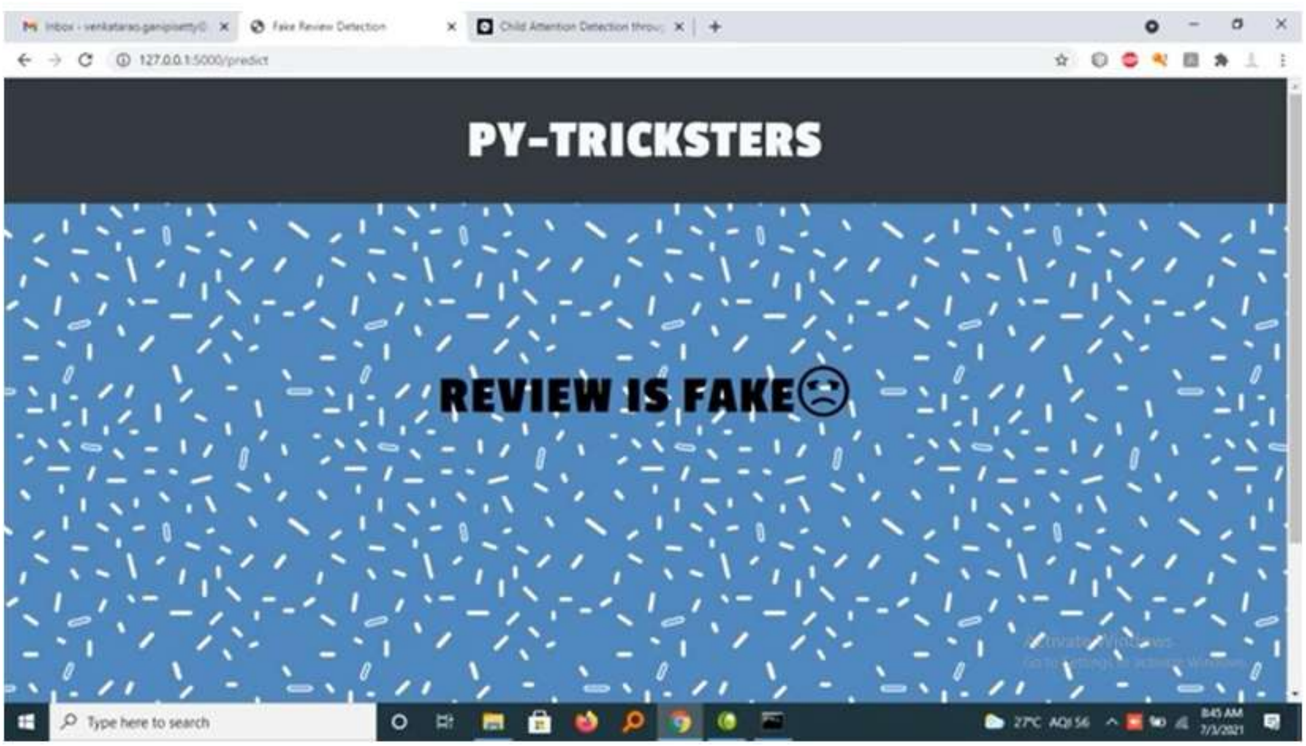

Fig-3: Output-1 Detecting review is fake or not

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 4, April : 2023

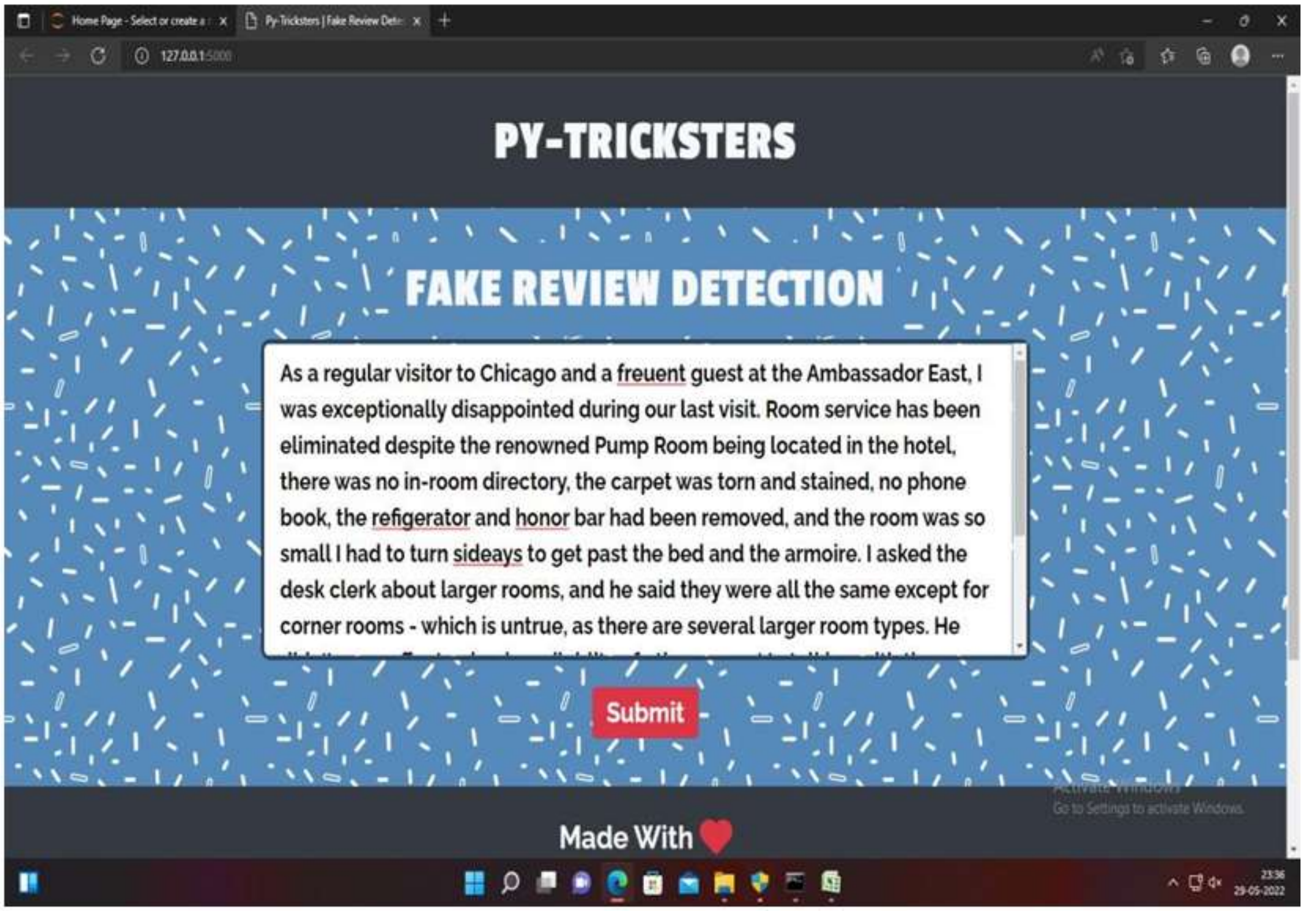

Fig-4: Input-2 Enter the review for detection

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 4, April : 2023

Fig-5: Output-2 Detecting review is fake or not

6 CONCLUSION AND FUTURE WORK

This research provided a thorough analysis of the most significant papers on machine learning-based fake review identification that have been published to date. First, we looked into the feature extraction strategies that other scholars have employed.

Then, we described the datasets that were already available and how they were created. Then, using summary tables, we described some conventional machine learning models and neural network models used for false review detection. By increasing feature extraction and classifier construction, traditional statistical machine learning raises the performance of text classication models. Deep learning, on the other hand, enhances the presentation learning method, the structure of the algorithm, and new knowledge to improve performance. We also offered a comparison of a few transformers and neural network model-based deep learning techniques that haven't been applied to the identification of false reviews. Results indicated that Roberta had the best accuracy across both datasets. Furthermore, Roberta's effectiveness in spotting bogus reviews was demonstrated by its recall, precision, and F1 score. We concluded by summarising the current research gaps and potential future directions to produce reliable results in this field.

We can draw the conclusion that the majority of previous studies used supervised machine learning to identify bogus reviews. To determine if a review is false or not, supervised machine learning requires a labelled dataset, which might be challenging to find in a fake review detection field. We noticed that the most often used datasets in the current studies are built based on a crowd sourcing framework due to the difficulties of getting labelled datasets. It is not advisable to evaluate machine learning techniques on these datasets because they do not depict the fake review in a practical application.

7 REFERENCES

1) E. I. Elmurngi and A.Gherbi, "Unfair Reviews Detection on Amazon Reviews using Sentiment Analysis with Supervised Learning Techniques," Journal of Computer Science, vol. 14, no. 5, pp. 714–726, June 2018.

2) J. Leskovec, "WebData Amazon reviews," [Online]. Available: http://snap.stanford.edu/data/web-Amazon-links.html [Accessed: October 2018].

3) J. Li, M. Ott, C. Cardie and E. Hovy, "Towards a General Rule for Identifying Deceptive Opinion Spam," in Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics, Baltimore, MD, USA, vol. 1, no. 11, pp. 1566-1576,November 2014.

4) N. O'Brien, "Machine Learning for Detection of Fake News," [Online]. Available: https://dspace.mit.edu/bitstream/handle/1721.1/119727/1078649610-MIT.pdf [Accessed: November 2018].J. C. S. Reis, A. Correia, F.

5) Murai, A. Veloso, and F. Benevenuto, "Supervised Learningfor Fake News Detection," IEEE Intelligent Systems, vol. 34, no. 2, pp. 76-81, May 2019.

6) B. Wagh, J. V. Shinde and P. A. Kale, "A Twitter Sentiment Analysis Using NLTK and Machine Learning Techniques," International Journal of Emerging Research in Management and Technology, vol. 6, no. 12, pp. 37-44, December 2017.

7) McCallum and K. Nigam, "A Comparison of Event Models for Naive Bayes Text Classification," in Proceedings of AAAI-98 Workshop on Learning for Text Categorization, Pittsburgh, PA, USA, vol. 752, no. 1, pp. 41-48, July 1998.

8) Liu and M. Hu, "Opinion Mining, Sentiment Analysis and Opinion Spam Detection," [Online]. Available: https://www.cs.uic.edu/~liub/FBS/sentiment-analysis.html#lexicon [Accessed: January 2019].

9) Hill, "10 Secrets to Uncovering which Online Reviews are Fake," [Online]. Available: https://www.marketwatch.com/story/10-secrets-to-uncovering-which-online-reviews-are-fake-2018-09-21 [Accessed: March 2019].

10) J. Novak, "List archive Emojis," [Online]. Available: https://li.st/jesseno/positive-negative-and-neutral-emojis-6EGfnd2QhBsa3t6Gp0FRP9 [Accessed: June 2019].

11) P. K. Novak, J. Smailović, B. Sluban and I. Mozeti, "Sentiment of Emojis," Journal of Computation and Language, vol.10, no. 12, pp. 1-4, December 2015.

12) P. K. Novak, "Emoji Sentiment Ranking," [Online].Available:http://kt.ijs.si/data/Emoji_sentiment_ranking/ [Accessed: July 2019]

13) E. F. Cardoso, R. M. Silva, and T. A. Almeida, ``Towards automatic ltering of fake reviews,'' Neurocomputing, vol. 309, pp. 106116, Oct. 2018.

14) L. Da Xu,W. He, and S. Li, ``Internet of Things in industries: A survey,'' IEEE Trans. Ind. Informat., vol. 10, no. 4, pp. 22332243, Nov. 2014.

15) Y. Ren and Y. Zhang, ``Deceptive opinion spam detection using neural network,'' in Proc. 26th Int. Conf. Comput.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 4, April : 2023

Linguistics: Tech. Papers (COLING), 2016, pp. 140150.

16) N. Jindal and B. Liu, ``Opinion spam and analysis,'' in Proc. Int. Conf. Web Search Web Data Mining (WSDM), 2008, pp. 219230.

17) Heydari, M. Tavakoli, and N. Salim, ``Detection of fake opinions using time series,'' Expert Syst. Appl., vol. 58, pp. 8392, Oct. 2016.

18) L. Li,W. Ren, B. Qin, and T. Liu, ``Learning document representation for deceptive opinion spam detection,'' in Chinese Computational Linguistics and Natural Language Processing Based on Naturally Annotated Big Data. Nanjing, China: Springer, 2015, pp. 393404.

19) H. Aghakhani, A. Machiry, S. Nilizadeh, C. Kruegel, and G. Vigna, ``Detecting deceptive reviews using generative adversarial networks,'' in Proc. IEEE Secur. Privacy Workshops (SPW), May 2018, pp. 8995.

20) Mukherjee, V. Venkataraman, B. Liu, and N. Glance, ``Fake review detection: Classication and analysis of real and pseudo reviews,'' Univ. Illinois Chicago, Chicago, IL, USA, Tech. Rep. UIC-CS-03- 2013, 2013.

21) R. Yafeng, J. Donghong, Z. Hongbin, and Y. Lan, ``Deceptive reviews detection based on positive and unlabeled learning,'' J. Comput. Res. Develop., vol. 52, no. 3, p. 639, 2015.

22) R. Y. K. Lau, S. Y. Liao, R. C.-W. Kwok, K. Xu, Y. Xia, and Y. Li, ``Text mining and probabilistic language modeling for online review spam detection,'' ACM Trans. Manage. Inf. Syst., vol. 2, no. 4, pp. 130, Dec. 2011.