

# Detecting and Analysing Fake Opinions Using Artificial Intelligence Algorithms

Mosleh Hmoud Al-Adhaileh<sup>1</sup> and Fawaz Waselallah Alsaade<sup>2,\*</sup>

<sup>1</sup>Deanship of E-learning and distance education, King Faisal University, Al-Ahsa, Saudi Arabia

<sup>2</sup>College of Computer Science and Information Technology, King Faisal University, Al-Ahsa, Saudi Arabia

\*Corresponding Author: Fawaz Waselellah Alsaade. Email: falsaade@kfu.edu.sa

Received: 27 June 2021; Accepted: 28 July 2021

**Abstract:** In e-commerce and on social media, identifying fake opinions has become a tremendous challenge. Such opinions are widely generated on the internet by fake viewers, also called fraudsters. They write deceptive reviews that purport to reflect actual user experience either to promote some products or to defame others. They also target the reputations of e-businesses. Their aim is to mislead customers to make a wrong purchase decision by selecting undesired products. Such reviewers are often paid by rival e-business companies to compose positive reviews of their products and/or negative reviews of other companies' products. The main objective of this paper is to detect, analyze and calculate the difference between fake and truthful product reviews. To do this, the methodology has planned to have seven phases: reviewing online products, analyzing features through linguistic enquiry and word count (LIWC), preprocessing the data to clean and normalize them, embedding words (Word2Vec) and analyzing performance using artificial deep-learning algorithms for classifying fake and truthful reviews. Two deep-learning neural network models have been evaluated based on standard Yelp product reviews. These models are bidirectional long-short term memory (BiLSTM) and convolutional neural network (CNN). The results from comparing the performance of the two models showed that the BiLSTM model provided higher accuracy for detecting fake reviews than the CNN model.

**Keywords:** Fake opinions; product reviews; fraudulent; e-business; deep learning

## 1 Introduction

Fake opinions detection is a subfield of natural language processing (NLP) that aims to analyse deceptive products reviews on e-business platforms. A deceptive products review are fabricated opinions, intentionally written by fraudulent people to seem trustworthy [1]. Consumers and e-companies often use online product reviews for procurement and organizational decisions because they include a wealth of knowledge. This knowledge is also a valuable resource for public opinion; it can affect resolutions over a wide spectrum of everyday and professional pursuits. Positive reviews can lead to significant financial gains, improve a company's reputation, while negative reviews can cause financial loss, and defame an e-business [2]. As a result, fraudsters have significant incentives to manipulate an opinion mining system by writing bogus reviews to support or disparage certain products or businesses [3]. These fraudsters are



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

also known as opinion scammers, and their behaviours are referred to as opinion spoofing or opinion spamming. Spam reviews have become a pervasive problem in e-commerce, with numerous high-profile incidents published in the media [4]. Developing intelligent opinion mining system is essential for any e-commerce company to identify an indication of deceptive review and help consumers to avoid bad choices during purchasing process [5]. Electronics products are the most frequently evaluated on online e-commerce websites [6], Such products often involve an important investment, and they are considered precious and valuable items. Therefore, they are frequently researched on e-commerce websites. As noted in [7], the decision to buy electronics product can be depended on the choices that are motivated by internet reviews, determining 24% of products obtained under this category. This is followed by word-of-mouth information as the second most famous source of opinion after using search engines to research these products. Moreover, shoppers commonly choose to buy electronics products because such products are changing continually by introducing smart products and improving existing ones [8]. Consequently, people frequently rely on online reviews to evade making faulty buying decisions [9].

Fake opinions include dishonest or inaccurate information. They are used mainly to misinform consumers so they make wrong purchase decisions, thus affecting the revenues for products. Spam product reviews are three types: Deceitful (fake) reviews, Reviews of a brand only and Non-reviews. 1) Deceitful (fake) reviews of products that have been written purposefully to mislead readers or systems of opinion mining. They include undeserving positive reviews to promote the online trade of specific products and negative reviews to defame worthy products. This type of spam product review is called hyperactive spam products reviews. 2) Reviews of a brand only: these opinions target the manufacturer brands instead of the product itself. 3) Non-reviews, which have two sub-kinds: (a) announcements and (b) unrelated reviews that contain no opinions, such as interrogations, responses or undefined text [10].

In this analysis, we focus on the linguistics of the review and behavioural of the reviewer features to identify, analyse and classify fake and truthful product reviews. The objective of this study is to demonstrate that word embeddings that incorporate artificial deep learning algorithms are an encouraging method for identifying fake opinions about online electronic consumer products. The remaining sections of this paper are Section 2 introduce related work, Section 3 gives details of the framework for the proposed methodology, and Section 4 presents wordcloud for most frequency words in the dataset, Section 5 section explore a comparative analysis of the obtained results and Section 6 gives conclusions and future work.

## 2 Related Work

Identifying fake reviews has become a prevalent research area in the last decade. Many researches have proposed different methodologies for identifying fake/spam/deceptive reviews because of their substantial influence on customers and e-commerce businesses. Different methods to do this have been presented, depending on the forms of data being analysed. These have included annotated data with supervised learning techniques (classification), unmarked data with unsupervised learning (clustering) and partially annotated data based on semi-supervised learning. The first research to analyse opinion spam was reported by Jindal et al. [11], who used duplicate or nearly duplicate reviews from the Amazon website to conduct their experimental work. Based on the review, reviewer and product features, they applied a supervised logistic regression algorithm to classify the reviews as fake or trusted reviews, and they attained 78% accuracy.

Ahmed et al. [12] focused on N-grams as features of the contents of reviews and a linear support vector machine (L-SVM) to detect fake reviews. They used a benchmark dataset, which contained fake hotel reviews collected from the [TripAdvisor.com](https://www.tripadvisor.com) website. For feature transformation and representation, they used Term Frequency-Inverse Document frequency (TFIDF) method. Their experimental result was an

F1 score of 90%. Ott et al. [13] tried to delineate a general rule for classifying unreliable reviews. In their approach, they used multi-label class cross-domain datasets that contained reviews of 800 hotels assembled from the Amazon Mechanical Turk (AMT) website and 400 deceptive doctor reviews collected from experts' domains. Concerning features, they adopted unigram, LIWC and Part of Speech (POS) in their experiments. For classification, they applied an SVM and the sparse additive generative (SAGE) model, which included an aggregate of the generalized additive topic models. Their results were 78% and 81% accuracy, respectively.

Savage et al. [14] proposed a method for considering and spotting opinion spammers based on the abnormality of the rating of products that was given by specific persons. The authors focused on variations among the deviated ratings and the evaluation ratings of a large number of opinions of honest reviewers. They computed the spamicity and honesty for each reviewer using binomial regression to detect reviewers who had an abnormal attribution of ratings that deviated from public opinion. Narayan et al. [15] applied PU-learning, using several models. In this regard, the authors applied six dissimilar classifiers to discover fake reviews. The classifiers were: decision tree, random forest, naive Bayes, logistic regression, SVM and k-nearest neighbour. They found that the logistic regression model classifier obtained the best performance of the six diverse algorithms. Rosso et al. [16] evaluated classical PU-learning in addition to adjusted PU-learning. Applying a modified PU-learning method, the authors found it feasible to discover fewer samples from the unmarked dataset. In each phase, only unseen negative cases (which are produced by yielding earlier iterations) were recognized, and the classifier was just involved in novel negative samples. Therefore, in every repetition, negative instances were diminished and ultimate instances were accurately classified as fraudulent or legitimate. They adopted Naïve Bayes (NB) and SVM classifiers with N-gram features. Lau et al. [17] proposed a probabilistic language model for detecting fraudulent reviews, adopting an Linear Discriminant Analysis (LDA) technique called the Topic Spam model, which classified reviews by computing the probability of words frequency in spam and non-spam contents.

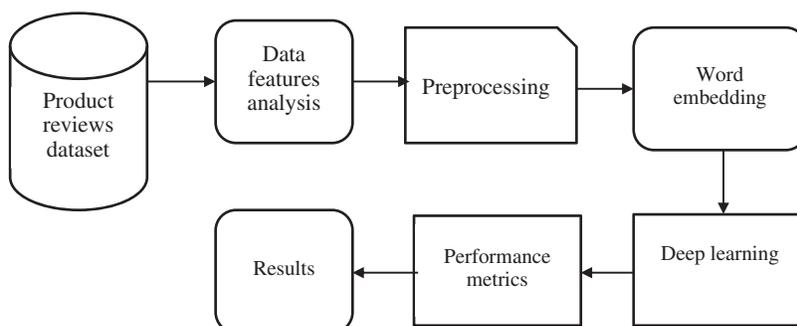
Based on deep learning, Goswami et al. [18] proposed a model to examine the effects of behavioural associations of reviewers to identify fake online reviews. The dataset they used was from the Yelp website. The data were collected and preprocessed for analysis. Further, they extracted social and behavioural indications of customers and fed them to a backpropagation neural network in order to categorize reviews as trusted or deceptive. They had a detection rate that was 95% accurate. Ren et al. [19] combined two neural network models that were constructed to detect opinion spam using in-domain and cross-domains reviews: the gated recurrent neural network (GRNN) and the convolutional neural network (GRNN-CNN). These datasets were related to doctors, restaurants, and hotel domains with corresponding sizes of 432, 720 and 1280 reviews. By merging all these reviews, they implemented their approach to classify reviews as truthful or fake. Their experimental results attained an accuracy of 83%.

Zeng et al. [20] used an RNN-BiLSTM technique to recognise counterfeit reviews. They split the review contents into three portions: first sentences, middle sentences, and last sentences. They found that the middle sentences had more fake clues than other portions. Their results had a value of 85% accuracy. In [21] analysis, the authors focused on singleton review junk recognition. The authors perceived that the configuration of truthful reviewers' appearance was stable and unconnected to their evaluation pattern, while spammers demonstrated contradictory actions. Those actions could be distinguishable to complete unusually connected temporal indications. However, the single review junk discovery issue was deemed tricky because of irregularly linked pattern recognition. They reached 75.86% and 61.11% accuracy, based on recall and precision, respectively. Alsubari et al. [22] proposed machine learning based models using supervised learning techniques such as , AdaBoost, Random Forest and decision tree. Further, the authors used the Linguistic inquiry and word count (LIWC) tool to extract and calculate deep linguistic features from products reviews to distinguish between truthful and fake reviews. Their models were

evaluated based on standard benchmark Yelp products reviews dataset. For selecting important set of features, they used an Information Gain (IG). It was observed that the Ada Boost model provided the best performance for detecting fake reviews, reaching 97% accuracy.

### 3 Materials and Methods

This section presents the methodology used in this study. This methodology has seven phases: product reviews dataset, analysing features, preprocessing, word embedding, CNN/BiLSTM, performance metrics and results. The below Fig. 1 depicts the proposed methodology.



**Figure 1:** Proposed methodology

#### 3.1 Product Reviews Dataset

The demand for e-commerce has grown rapidly in recent years because of the ease of internet access in addition to innovative technologies. Various factors represent the fame of e-commerce businesses as well as their standings. They include credibility, product quality, and customer recommendation systems. Product reviews are opinions about products written by shoppers or customers. They are considered the most important factors for improving the credibility, standards and assessment of an e-commerce store. Using product reviews, owners of e-commerce businesses can detect and identify imperfections of their products and analyse the feelings and satisfaction of consumers [23]. The dataset used in this study comprised benchmark Yelp product reviews that were collected [24]. It comprised 9456 fake electronic product reviews gathered from four American cities: Los Angeles, Miami, New York and San Francisco. This dataset was labelled using a labelling algorithm that is utilized to filter fake reviews on the [Yelp.com](https://www.yelp.com) website [25].

#### 3.2 Analyzing the Features of the Dataset

To analyse and find the difference between fake and truthful reviews, all reviews of the used dataset were inserted into the linguistic inquiry and word count (LIWC) dictionary [26], which is mostly used in text mining for natural language processing (NLP) tasks. Each review structure in the dataset had meta-features such as rating value, review text, reviewer name, product ID and class label. Furthermore, we extracted additional significant linguistic features from the review content using the LIWC dictionary: authenticity of the review text, analytical thinking of the reviewer, word count, negative words, positive words, and negation words. Tab. 1 shows the statistical averages of the features set for 4790 fake reviews and 4666 truthful reviews.

**Table 1:** The average values for the features set of a dataset

Feature name	Fake reviews			Truthful reviews		
	Mean	SD	MAX	Mean	SD	MAX
Authenticity	32.95	24.18	99	74.84	20.89	99
Analytical thinking	66.51	22.0	99	44.59	22.61	98
Rating value	3.8	1.65	5	3.4	1.79	5
Word count	124	117	1010	103	102	994
Personal pronouns	9.37	4.28	26	12	3.80	31
Negation words	1.83	1.30	14	2.42	1.76	23
Positive words	5.5	5.19	111	4.0	3.1	33
Negative words	1.94	2.0	51	1.96	1.59	25

### 3.3 Preprocessing

Before carrying out the transformation and vectorization of the sentences of the reviews, preprocessing steps were used to clean the data and remove noise. The goal of text preprocessing is to convert the texts of the reviews to a form that deep learning algorithms can understand and analyse. The preprocessing steps are as follows:

- a) Removing punctuation: deleting punctuation marks from the reviews.
- b) Removing stopwords: This process cleans articles from the text; for example, ‘the’, ‘a’, ‘an’ words are removed from text.
- c) Stripping useless words and characters from the dataset.
- d) Word stemming: converting each word of a sentence into its root; for instance, ‘undesired’ becomes ‘desire’
- e) Tokenizing: splitting whole sentences in the text into separate words, keywords, phrases, and pieces of tokens.
- f) Padding sequences: using deep learning neural networks to ensure that the input data have equal sequence length. However, we implemented a pre-padding method to add zeros to the beginnings of the vector representation.

### 3.4 Word Embedding's

Word2Vec is a technique that is used for attaining word embedding representation of text [27]. It can obtain and learn the semantics between word representation vectors in textual sentences [28]. This method was created by Toke et al. [29] in 2013 to map words of a text based on their associations and meanings. The Word2Vec has two types of approaches: continuous bag-of-word (CBOW) and skip-gram. By using this technique, the relationship between words is calculated through a cosine similarity function. Furthermore, it map each word of a review text into real-valued vector named word embedding.

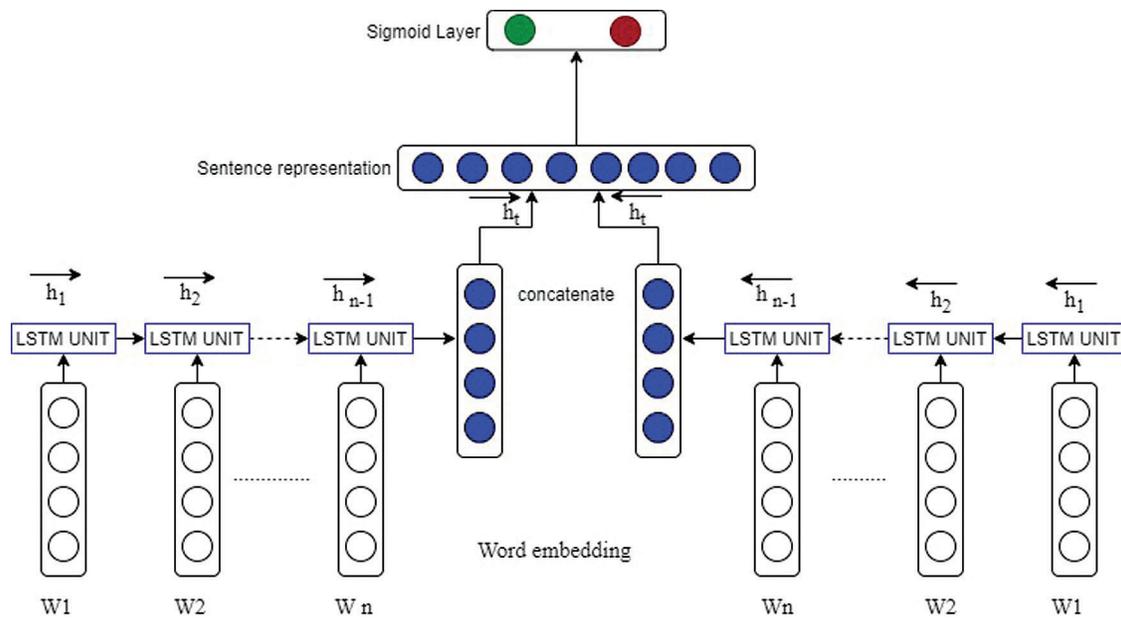
### 3.5 Classification

In this section, deep learning neural network models are presented for analysing, detecting and classifying fake reviews based on n-gram features of review text into a fake or truthful review. These are

the convolutional neural network (CNN) model and the bidirectional long short-term memory (BiLSTM) model.

### 3.5.1 Bidirectional Long Short-Term Memory (BiLSTM) Model

Recurrent neural networks (RNN) are types of deep learning techniques that have a repetitive deep hidden nature, and the hidden state is stimulated by the earlier states at a specific time. Hence, RNNs have the proficiency to model the contextual data effectively and process the changeable-length series of input data. LSTM is a form of the RNN structure, and it has recently enhanced the design of RNNs. LSTM solved the issue of vanishing gradient points by substituting the self-linked hidden layers with memory units. The memory units are utilized for storing long-range information of input data when processing [30]. In the case of the LSTM, the process of data handling can take place in only one direction, which means in a forward direction. This disregards backward production, so it reduces the performance of the system. To mitigate this shortcoming, the input data are handled in bidirectional recurrent neural networks by concatenating the production of forward and backward directions. Fig. 2 describes the structure of the BiLSTM model for text classification to detect fake reviews. In this model, the sentences process of embedding sentences' words was performed based on the Word2Vect method that is explained in the previous section.



**Figure 2:** Structure of the bidirectional long short term memory model

Every LSTM unit has four gates: input  $i_t$ , forget  $f_t$ , cell state  $c_t$  and output gate  $o_t$  which are formulated as follows [31].

$$i_t = \sigma(W_{ix}x_t + W_{ih}h_{t-1} + b_i) \quad (1)$$

$$f_t = \sigma(W_{fx}x_t + W_{fh}h_{t-1} + b_f) \quad (2)$$

$$o_t = \sigma(W_{ox}x_t + W_{oh}h_{t-1} + b_o) \quad (3)$$

$$c_t = f_t c_{t-1} + i_t * \tanh(W_{cx}x_t + W_{ch}h_{t-1} + b_c) \tag{4}$$

$$\vec{h}_t = o_t * \tanh(c_t) \tag{5}$$

$$\overleftarrow{h}_t = o_t * \tanh(c_t) \tag{6}$$

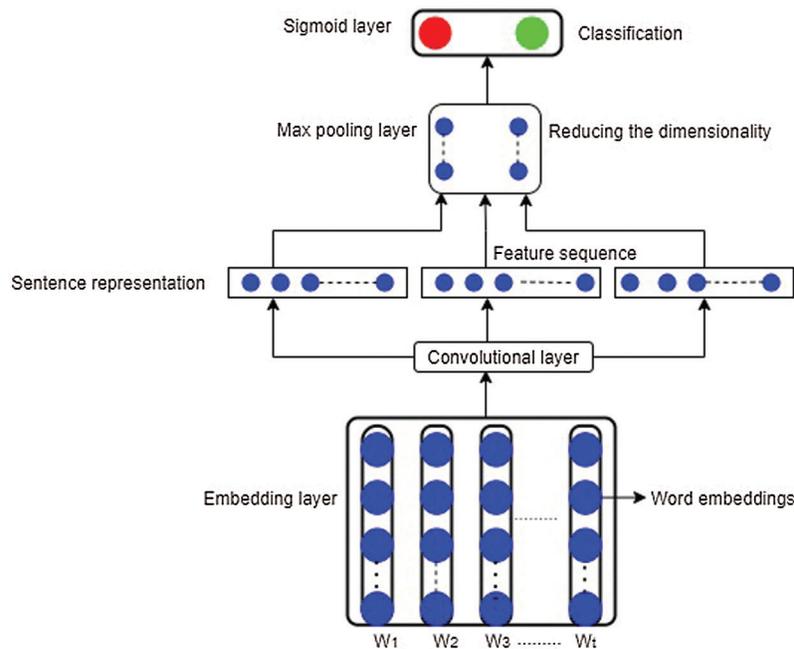
$$\tanh(x) = \frac{1 - e^{2x}}{1 + e^{2x}} \tag{7}$$

$$H_t = (\vec{h}_t : \overleftarrow{h}_t) \tag{8}$$

where  $\vec{h}_t$  states the production of the LSTM cell,  $H_t$  is the output of bidirectional concatenation of  $\vec{h}_t$  forward and  $\overleftarrow{h}_t$  backward LSTM layers at the current time state  $t$ .

### 3.5.2 The CNN Model

Convolutional neural networks (CNNs) are a kind of artificial forward neural network broadly applied for composing a semantic pattern [32], apprehending n-gram information in an automated manner. In this model, an embedding layer-based Word2Vec technique is organized through three parameters: vocabulary size, length of input sequence and embedding dimension. The vocabulary size signifies the most frequent words in the dataset, and it is set to 40,000 words. Word embeddings gives the dimensions vectors for each word of a review. A maximum input sequence length is the maximum length of the review text. The convolution layer leads an operation that uses filters on the input embeddings matrix  $E(w) \in R^{V \times D}$ , where  $V$  is vocabulary size, and  $D$  indicates the word dimensions. We set 100 filters with a window size of 3. The max pooling layer takes the input sequences from the convolutional layer, implements a downsampling process in addition to reducing spatial dimensionality for the given sequences, and selects a maximum value of feature sequences from the pool of each filter kernel. The CNN model has been used to learn word embeddings of sentences and forward their real-valued vectors to the classification layer. This is a sigmoid activation function layer applied to calculate the probability that each input sequence of the review text would be classified as fake or truthful. Fig. 3 illustrates the architecture of the CNN model.



**Figure 3:** Structure of the CNN model

## 4 Experimental Results

We performed several experiments on the standard Yelp dataset, which includes 9456 electronic product reviews. The dataset was divided into 70% as a training set, 10% as a validating set, and 20% as the testing set. This subsection presents the experimental results of the CNN- and BiLSTM-based deep learning models for detecting fake reviews using learning word embedding representations. Both models have been trained and validated five times with a batch size of 300 samples of review texts per time step. By comparing the performance of the CNN and BiLSTM techniques, the results showed that the BiLSTM model was more accurate than the CNN model, while the CNN model required less data processing time. Fig. 6 displays the classification results for both models.

### 4.1 Metrics for Model Performance

To evaluate the performance of the proposed CNN and BiLSTM models for detecting and predicting fake and truthful reviews, various performance metrics can be calculated from the confusion matrixes based on the rates of false-positive and false-negative items. Assessment metrics like specificity, precision, recall, F1-score, and accuracy were computed from two confusion matrixes, as shown in Figs. 4 and 5.

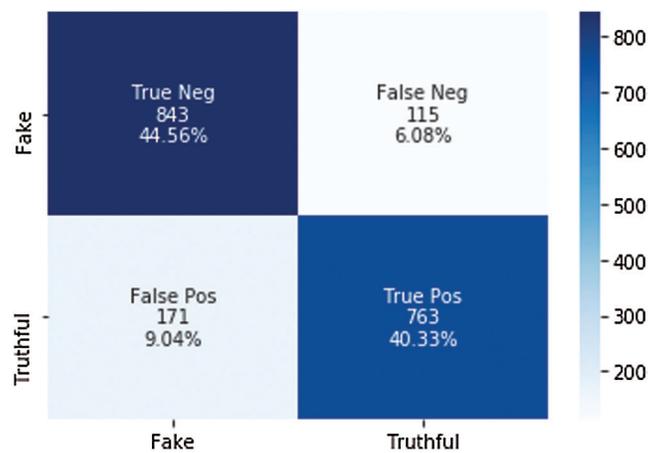


Figure 4: Confusion matrix for the BiLSTM model

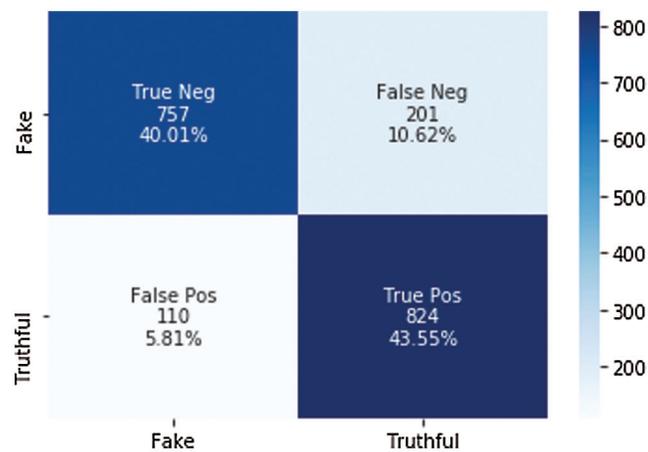
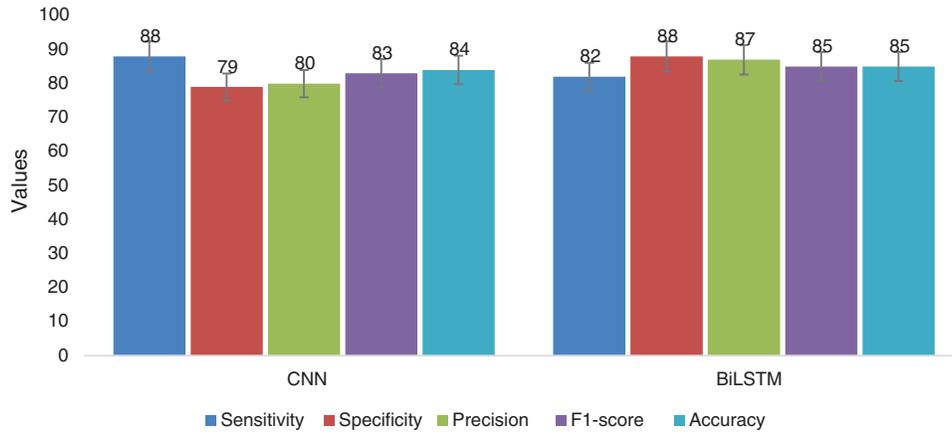


Figure 5: Confusion matrix for the CNN model



**Figure 6:** Classification results for the CNN and BiLSTM models

These evaluation metrics are formulated as follows.

$$Sensitivity = \frac{TP}{TP + FN} \times 100 \quad (9)$$

$$Specificity = \frac{TN}{TN + FP} \times 100 \quad (10)$$

$$recision = \frac{TP}{TP + FP} \times 100 \quad (11)$$

$$F1 - score = 2 * \frac{precision \times sensitivity}{precision + sensitivity} \times 100 \quad (12)$$

$$Accuracy = \frac{TP + TN}{FP + FN + TP + TN} \times 100 \quad (13)$$

where TP designates the number of reviews that were properly classified as truthful reviews, FP indicates the number of reviews that were incorrectly identified as fake reviews, TN represents the total number of reviews that were successfully classified as fake reviews, and FN is the number of reviews properly classified as truthful reviews.

In Fig. 7a shows Y-axis accuracy of a BiLSTM model to fake reviews detection and X-axis represents the epochs that are number of iterations of training and testing the model on the used dataset whereas in the Fig. 7b, illustrates Y-axis the training model loss and X-axis gives the number of epochs.

In Fig. 8c shows Y-axis accuracy of a CNN model to fake reviews detection and X-axis represents the epochs that are number of iterations of training and testing the model on the used dataset whereas in the Fig. 8b shows displays Y-axis illustrates the training model loss and X-axis gives the number of epochs.

#### 4.2 Word Cloud

A wordCloud is technique used widely to visualize a textual data and allows the researchers to observe in a single scan the words that are having the maximum occurrence in a given body of text. Fig. 9 visualizes the most repeated words in the used dataset.



## 5 Comparative Analysis

This section presents comparative analysis between the results obtained by the proposed deep learning based models and machine learning based model presented. Furthermore, it provides comparing between transformation methods adopted to map the words of text into numerical form are introduced. In case of TF-IDF, the word in the review text is converted into single column vector whereas in Word2Vec method, it is transformed into N-dimensional vectors. Tab. 2 compares the proposed models with some existing work based on an accuracy metric and the same dataset used in this study.

**Table 2:** Comparison of the obtained results with existing work based on accuracy metric

Paper ID	Dataset used	Feature representations	Technique	Accuracy (%)
Barbado et al. [33]	Yelp product reviews	TF-IDF	Ada boost	82
This study	Yelp product reviews	Word2Vec	CNN	84
			BiLSTM	85

## 6 Conclusions

Fake reviews are deceptive opinions posted by fraudulent reviewers on the websites of e-business companies to mislead customers. The aim is to have the customers make wrong purchases by selecting an inferior product. In this paper, we used standard Yelp product reviews to conduct empirical tests of the performance of two neural network models, CNN and BiLSTM, for identifying fake opinions. Based on the learning word embeddings of the text of reviews, the proposed models were used to classify those product reviews as fake or truthful. By comparing the results obtained from several experiments, we found that the BiLSTM model provides higher performance than the CNN model. In text classification tasks, neural networks can appropriately catch global semantic information over sentence vectors. As a result, deep learning-based models outperform the baseline models in terms of accuracy. The experimental results also showed that the CNN model was superior to the BiLSTM model in data processing time. In future work, we will attempt to develop a hybrid neural networks-based model using behavioural and linguistic features for detecting fake reviews.

**Acknowledgement:** The authors extend their appreciation to the Deanship of Scientific Research at King Faisal University for funding this research work through the project number No. 206017.

**Funding Statement:** This research and the APC were funded by the Deanship of Scientific Research at King Faisal University for the financial support under grant No. 216017.

**Conflicts of Interest:** The authors declare that they have no conflicts of interest to report regarding the present study.

## References

- [1] Y. Ren and D. Ji, "Neural networks for deceptive opinion spam detection: An empirical study," *Information Sciences*, vol. 385, no. 2, pp. 213–224, 2017.
- [2] N. Hussain, H. T. Mirza, I. Hussain, F. Iqbal and I. Memon, "Spam review detection using the linguistic and spammer behavioral methods," *IEEE Access*, vol. 8, no. 4, pp. 53801–53816, 2020.
- [3] D. U. Vidanagama, T. P. Silva and A. S. Karunananda, "Deceptive consumer review detection: A survey," *Artificial Intelligence Review*, vol. 53, no. 2, pp. 1323–1352, 2020.

- [4] L. Xiang, G. Guo, Q. Li, C. Zhu, J. Chen *et al.*, “Spam detection in reviews using lstm-based multi-entity temporal features,” *Intelligent Automation & Soft Computing*, vol. 26, no. 6, pp. 1375–1390, 2020.
- [5] L. Xiang, G. Guo, J. Yu, V. S. Sheng and P. Yang, “A convolutional neural network-based linguistic steganalysis for synonym substitution steganography,” *Mathematical Biosciences and Engineering*, vol. 17, no. 2, pp. 1041–1058, 2020.
- [6] Y. Y. Yolanda, E. Chan and W. T. Ngai, “Conceptualising electronic word of mouth activity: An input-process-output perspective,” *Marketing Intelligence & Planning*, vol. 29, no. 5, pp. 488–516, 2011.
- [7] C. Riegner, “Word of mouth on the web: The impact of web 2.0 on consumer purchase decisions,” *Journal of Advertising Research*, vol. 47, no. 4, pp. 436–447, 2007.
- [8] U. Chakraborty and S. Bhat, “The effects of credible online reviews on brand equity dimensions and its consequence on consumer behavior,” *Journal of Promotion Management*, vol. 24, no. 1, pp. 57–82, 2018.
- [9] M. Z. Asghar, A. Ullah, S. Ahmad and A. Khan, “Opinion spam detection framework using hybrid classification scheme,” *Soft Computing*, vol. 24, no. 5, pp. 3475–3498, 2020.
- [10] P. Do-Hyung Park and S. Kim, “The effects of consumer knowledge on message processing of electronic word-of-mouth via online consumer reviews,” *Electronic Commerce Research and Applications*, vol. 7, no. 4, pp. 399–410, 2008.
- [11] N. Jindal and B. Liu, “Opinion spam and analysis,” in *Proc. of the 2008 Int. Conf. on Web Search and Data Mining*, Palo Alto, California, USA, pp. 219–230, 2008.
- [12] H. Ahmed, I. Traore and S. Saad, “Detecting opinion spams and fake news using text classification,” *Security and Privacy*, vol. 1, no. 1, pp. 1–15, 2018.
- [13] M. Ott, Y. Choi, C. Cardie and J. T. Hancock, “Finding deceptive opinion spam by any stretch of the imagination,” in *Proc. of the 49th Annual Meeting of the Association for Computational Linguistics*, Portland, Oregon, pp. 309–319, 2011.
- [14] D. Savage, X. Zhang, X. Yu, P. Chou and Q. Wang, “Detection of opinion spam based on anomalous rating deviation,” *Expert Systems with Applications*, vol. 42, no. 22, pp. 8650–8657, 2015.
- [15] R. Narayan, J. Rout and S. Jena, “Review spam detection using semisupervised technique,” *Progress in Intelligent Computing Techniques: Theory, Practice, and Applications*, vol. 15, no. 2, pp. 281–286, 2018.
- [16] P. Rosso, D. Cabrera and M. Gomez, “Detecting positive and negative deceptive opinions using PU-learning,” *Information Processing and Management*, vol. 51, no. 4, pp. 1–11, 2014.
- [17] R. Y. Lau, S. Y. Liao, R. C. Kwok, K. Xu, Y. Xia *et al.*, “Text mining and probabilistic language modeling for online review spam detection,” *ACM Transaction, Management*, vol. 2, no. 5, pp. 233–245, 2011.
- [18] K. Goswami, Y. Park and C. Song, “Impact of reviewer social interaction on online consumer review fraud detection,” *Journal of Big Data*, vol. 4, no. 1, pp. 1–19, 2017.
- [19] Y. Ren and D. Ji, “Neural networks for deceptive opinion spam detection: An empirical study,” *Infection Science*, vol. 385, no. 8, pp. 213–224, 2017.
- [20] Z. Y. Zeng, J. J. Lin, M. S. Chen, M. H. Chen, M. H. Lan *et al.*, “A review structure based ensemble model for deceptive review spam,” *Information-an International Interdisciplinary Journal*, vol. 10, no. 7, pp. 243, 2019.
- [21] S. Xie, G. Wang, S. Lin and P. S. Yu, “Review spam detection via temporal pattern discovery,” in *Proc. of the 18th ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining, KDD '12*, New York, NY, USA, pp. 823–831, 2012.
- [22] S. N. Alsubari, M. B. Shelke and S. N. Deshmukh, “Fake reviews identification based on deep computational linguistic features,” *International Journal of Advanced Science and Technology*, vol. 29, no. 8, pp. 3846–3856, 2020.
- [23] Y. Wang, B. Liu, H. Wu, S. Zhao, Z. Cai *et al.*, “An opinion spam detection method based on multi-filters convolutional neural network,” *Computers, Materials & Continua*, vol. 65, no. 1, pp. 355–367, 2020.
- [24] Z. Y. Xiong, Q. Q. Shen, Y. J. Wang and C. Y. Zhu, “Paragraph vector representation based on word to vector and CNN learning,” *Computers, Materials & Continua*, vol. 55, no. 2, pp. 213–227, 2018.

- [25] T. H. M. Alrasheed, M. H. Al-Adaileh, A. A. Alqarni and M. Y. Alzahrani, "Deep learning and holt-trend algorithms for predicting covid-19 pandemic," *Computers, Materials & Continua*, vol. 67, no. 2, pp. 2141–2160, 2021.
- [26] M. L. Newman, J. W. Pennebaker, D. S. Berry and J. M. Richards, "Lying words: Predicting deception from linguistic styles," *Personality and Social Psychology Bulletin*, vol. 29, no. 5, pp. 665–675, 2003.
- [27] S. N. Alsubari, S. N. Deshmukh, M. H. Al-Adhaileh, F. W. Alsaade and T. H. Aldhyani, "Development of integrated neural network model for identification of fake reviews in e-commerce using multidomain datasets," *Applied Bionics and Biomechanics*, vol. 2021, no. 4, pp. 1–11, 2021.
- [28] L. Akoglu, R. Chandy and C. Faloutsos, "Opinion fraud detection in online reviews by network effects," in *Proc. of the 7th Int. Conf. on Weblogs and Social, Media*, Cambridge, Massachusetts, USA, pp. 2–11, 2013.
- [29] P. S. Toke, R. Mutha, O. Naidu and J. Kulkarni, "Enhancing text mining using side information," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 5, no. 5, pp. 793–797, 2016.
- [30] L. Gang and J. Guo, "Bidirectional LSTM with attention mechanism and convolutional layer for text classification," *Neurocomputing*, vol. 337, no. 7, pp. 325–338, 2019.
- [31] M. Luca and G. Zervas, "Fake it till you make it: Reputation, competition, and yelp review fraud," *Management Science*, vol. 62, no. 12, pp. 3412–3427, 2016.
- [32] H. Y. Li, Z. Y. Chen, A. Mukherjee, B. Liu and J. D. Shao, "Analyzing and detecting opinion spam on a large-scale dataset via temporal and spatial patterns," in *Proc. Conf. on Web and Social, Media*, Oxford, UK, pp. 634–637, 2015.
- [33] R. Barbado, O. Araque and C. A. Iglesias, "A framework for fake review detection in online consumer electronics retailers," *Information Processing & Management*, vol. 56, no. 4, pp. 1234–1244, 2019.