Fake Review Detection Using Deep Learning

Sangeetha S, Sangeetha B, Raja Kumar I, Shevannth R, Krishna Prasath S, Mohammed Rafi M

PSG College of Technology, Coimbatore, Tamil Nadu

Corresponding author: Sangeetha S, Email: vns.sangeetha@gmail.com

In recent times online shopping has evolved rapidly, but finding a quality product in such a complex network is not a simple task. Internet reviews helps users to find relevant items. But there is a high magnitude of fake internet reviews available online. So distinguishing fake and true reviews is an important task for both customers and suppliers. Because in the customer perspective they require a quality product and in the supplier perspective, they need to sell their product. Generally, positive reviews on a targeted product would increase its sales. To determine fake reviews this paper compares the reviews of several reviewers from Yelp dataset of restaurants and proposes a deep learning approach to detect fake reviews. The proposed deep learning model is tested with benchmarked datasets and results are evaluated. The experimental results shows improved performance compared to the existing deep learning models.

Keywords: Fake Review, Deep Learning, CNN, Bi Directional LSTM, Yelp Dataset.

2023. In Saroj Hiranwal & Garima Mathur (eds.), *Artificial Intelligence and Communication Technologies*, 655–667. Computing & Intelligent Systems, SCRS, India. https://doi.org/10.52458/978-81-955020-5-9-62

1 Introduction

Nowadays, when people make decision about services or products, reviews become the important source of their information. Based on the review given people buy things which increase the product sales. The positive review given to a product influence people to buy the product. Thus, reviews become very credible sources of information to most people in several online services. Since reviews are regarded authentic forms of giving genuine opinion regarding good or bad services, any attempt to manipulate those reviews by writing misleading or inauthentic content is considered as misleading action and such reviews are the fake information on a product. A fake review on a genuine product questions the credibility of the product. As a result, identifying fraudulent reviews has become an important and essential research topic.

2 Word Embedding Techniques

Word embedding are one of the popular representation of document language. It is able to acquire the context of a term in a text, semantic and syntactic likeness, and relation with other words.

2.1 GLoVe

It can be used to find association between words which are morpheme like zip-code and cities, etc. the unsu- pervised learning models are not constructive in associating homographs like word with the dissimilar meanings and identical spelling. GLoVe model forms vector for all the word, each vector have unique components, the words with matching components are very much similar.

The traditional one hot encoding provides a huge binary representation of the words. Processing such huge information is challenging and it does not cap- ture the semantic relationship between words. These drawbacks are addressed by glove model with a compact representation with a good understandability on semantic meaning.

3. Deep Learning Techniques

3.1 BiLSTM

BiLSTM stands for Bidirectional long short term memory is a popular deep learning technique. The BiLSTM level comprises of a collection of constantly connected memory blocks. These blocks are frequently allowed as a distinct form of the memory chips in an Information processing system. BiLSTM consist of inputs in forward as well as backward direction. The output can move to both front and back states. BiLSTM is very useful for classifying and predicting time series data

4. Related Works

S.N	Author	Title	Proposed System
1	Shaohua Jia, et al. [1]	Fake review detection based on LDA [2018]	In this paper, a multitude of machine learning algorithms have been used to identify which algorithm detects the most number of fake reviews posted on Yelp with better efficacy. After collecting the prerequisite datasets and extracting the features in it, a logistic regression model, SVM model and a Multi-layer perceptron model has been trained. It has been determined that the LDA model has produced the best results with an 81% accuracy rate.
2	Sanjay K. S and Dr. Ajit Danti [2]	Detection of fake opinions on online products using decision tree and information gain [2019]	Decision tree classification approach is used in this paper for differentiating genuine and fake reviews in shopping websites after taking six different parameters into consideration. The features of the datasets are extracted based on these parameters and a common template among the fake reviews has been obtained. The decision tree model is then trained with this template to ascertain the fake reviews.
3	Jane Crystal Rodrigues [3]	Machine and Deep Learning Techniques for detection of fake reviews: A Survey [2020]	In this paper six different machine learning algorithms have been reviewed for the given problem statement. It has been concluded that the sentiment analysis method

			provides one of the most accurate results when it comes to sorting spam reviews in websites. The paper combines three of the best result producing models namely, GRNN, BI-LSTM and LSTM.
4	Ahmed M. Elmogy [4]	Fake reviews detection using supervised machine learning [2021]	This paper follows a similar approach to that of the LDA method. It focuses purely on the K-nearest neighbor classifier method. It uses seven different parameters to extract the features and reengineer them, after collecting the dataset. This approach has the error rate of exactly 3.80%, which is considerably minimalistic.
5	SP.Rajamohan, et al. [5]	Survey on online review spam detection techniques [2017]	In this paper, the fake review datasets are initially divided into three segments; untruthful reviews, non-reviews and brand targeted reviews. It uses seven different approaches to the test and has indefinitely proved that the hybrid approach using the Naïve Bayes method provides the best outcome.
6	Yin Shuqin [6]	Fake Reviews Detection Based on Text Feature and Behaviour Feature [2019]	The goal of this study is to develop a fake reviews classification model using an MPINPUL (Mixing Population and Individual Nature PU Learning) model based on various features. There are four steps to the MPINPUL model. To begin, a limited k-means algorithm is provided for calculating a negative trust example.

			Multiple representative samples for positive and negative are calculated using LDA and k-means. The recognition rate of the MPINPUL model is demonstrated by experimental findings on real data sets.
7	Rakibul Hassan and Md. Rabiul Islam [7]	Impact of Sentiment Analysis in Fake Online Review Detection [2021]	This paper proposes a sentiment analysis approach that can effectively distinguish between positive and negative emotional reviews. It depicts a sentiment distribution analysis for both bogus and genuine reviews. Using a hotel review dataset, the proposed sentiment model is also utilized to investigate the influence of probabilistic sentiment score in the identification of false online reviews.
8	Nidhi A. Patel [8]	Survey on Fake Review Detection using Machine Learning Techniques[2018]	This paper discusses supervised, unsupervised, and semi-supervised data mining strategies for detecting bogus reviews using various features. They have used different features in detail like linguistic features, behavioral and relational features. They compared different techniques to identify fake reviews. They have also discussed major challenges of fake review detection.
9	G. M. Shahariar [9]	Spam Review Detection Using Deep Learning[2017]	This paper proposes to detect misleading reviews. In order to achieve it both labeled and

			unlabeled data are used. Deep learning methods for spam review detection which includes CNN MLP, CNN and LSTM are proposed. Some conventional machine learning classifiers such as SVM, KNN and Nave Bayes are used to detect spam reviews, they have compared both deep learning classifiers and conventional models.
10	Yue Shang [10]	Detecting Fake Reviews Using Multidimensional Representations With Fine-Grained Aspects Plan [2020]	In this paper, the focus is on the attention-based multilevel interactive neural network model with aspect constraints that mines the multilevel implicit expression mode of reviews. It integrates four dimensions, fine-grained aspects, products, users, review texts and namely, into review representations. The relationships between users and products are modeled and these relationships are used as a regularization term to redefine the model's objective function. The experimental results from three public datasets show that the model that is proposed is superior to the state-of-the-art methods; thus showing the effectiveness and portability of the model.
11	Dyar Wahyuni and Arif Djunaidy [11]	Fake Review Detection from a Product Review Using Modified Method of Iterative Computation[2016]	In this paper association rule mining approaches, social relationship analysis approaches and multi-task learning approaches are used to detect deceptive reviews. They stated that the precision value of rule based ICF++ is

			higher than ICF, with the inferential language model which outperformed all the other methods in TREC-like experiment. The RWR algorithm for trust-based rating predictions outperformed the conventional Computation Framework (ICF) technique, with a good link between social relationships and computed trustworthiness scores.
12	Dhairya Patel [12]	Fake Review Detection using Opinion Mining [2018]	This paper proposes a seller's inner motive to sell the product and make profit might affect the genuinely made feedback. Such feedback which he/she provides to the customer and if the customer does not consider it while buying the product then it may not be useful.
13	Aishwarya Pendyala [13]	Fake Consumer Review Detection [2019]	In our day-to-day life, we come across many social media and web applications, android applications. For all of these developments, as well as to determine whether this product is a success or a failure, they must examine client feedback. While doing this some customers genuinely use this product and give their review. These are called as auth reviews or true reviews. Someone who might try to destruct the product may give some false and negative review which are classified as spam reviews. So, in order to predict that, machine learning algorithms

			like random forest, svm are used to analyse the fake and true reviews. By this sentiment analysis of data, the product manager/owner is able to trace the original reviews and they work on their product growth. Accuracy of the proposed algorithm is 80.524%
14	Jingdong Wang [14]	Fake Review Detection Based on Multiple Feature Fusion and Rolling Collaborative Training [2020]	They used multiple feature fusion and rolling collaborative training and the most accurate classifier is selected to classify the new reviews. Here svm and random forest shows great accuracy, the result for supervised algorithm is 81.04% , for semi- supervised it is 82.53 ,co- training is 82.73% and for co-training(multi-feature-fusion) it is 84.45%.
15	Jay Kumar [15]	Fake Review Detection Using Machine Learning Techniques [2020]	Positive ratings of a specific object may attract more customers and improve sales, while unfavorable evaluations may result in lower demand and sales. These fake/fraudulent reviews are produced with the intent of deceiving potential consumers in order to promote/hype or discredit their businesses. This paper aims to determine whether a review is genuine or not.

5. Proposed Approach

Fig. 1 depicts our proposed model for spam review identification, which is briefed in this section. Our proposed model consists of four phases. Data Acquisition and Data Preprocessing is the initial phase, In Data Acquisition phase, around 2,00,000 instances are taken. The data was unbalanced. So the dataset is balanced by Under-sampling. After Under-sampling there are around 95,000 samples, which is adequate for both training and testing. In pre-processing phase, the relevant and meaningful features are selected for preprocessing.

The preprocessing performed are Vectorization and Padding Sequences to make all the reviews to same length. Other preprocessing techniques like removing punctuations, stemming were not applied because these serves as one of the effective features in differentiating the reviews between fake and genuine.

Then an Embedding Matrix is created using pre-trained Glove840B.300d Embedding file. The pre-trained embedding matrix is used to represent the characteristic features of individual terms in 300 dimensions. The Embedding Matrix is fed into Embedding layer with trainable as "False" so that pretrained weights won't get updated during the training process. The second phase of our proposed approach is Model Development. In this phase Bidirectional LSTM and CNN are used for model creation. Both the models take input from the Embedding Layer and process it separately to produce its own results and then output is concatenated and fed into Dense layer with linear activation function 'relu' and a Dropout layer (which is used to reduce overfitting) followed by Dense Layer with 2 neurons. Finally sigmoid activation function is used to obtain the output.

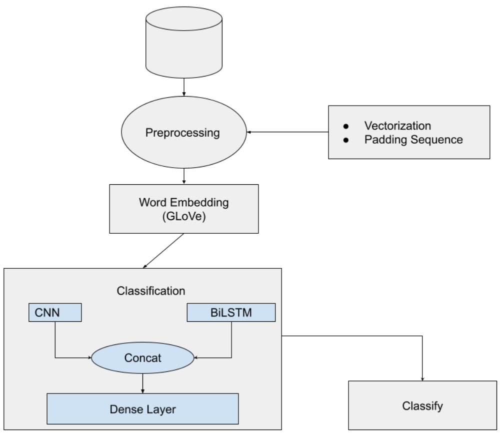

Fig 1. Proposed System.

Figure 2 illustrates the details of model architecture. A bunch of MaxPool- ing and GlobalMaxPooling lsince the model makes use ensures to capture all the overfit.

The next major part comes in Hyper-parameter tuning, we had played around with various Hyper-parameters such as number of neurons per layer, number of layers and batch size. If number of neurons and number of layers is too high, then the model will over-fit. Similarly, if it's too low, then the model won't be able to learn to resulting in High Bias and High Variance. In order to minimize the bias and variance, adequate layers with neurons should be designed, We find that our current configuration is optimal. The next hyper- parameter is batch size, Higher batch size leads to faster training and poor generalization. An optimal batch size is necessary for acceptable training time and good generalization. We find that batch size of 128 is optimal for yelp dataset.

The model is trained with 70,294 samples with 300 dimensions in 130 seconds. The "ADAM" optimizer is used for training with learning rate of 0.001(default).

6. Results

Figure 3 illustrates the comparison between other machine learning models like logistic regression, naïve Bayes, KNN, SVM and Random Forest and the highest accuracy other paper are able to achieve is 86.9% in SVM.

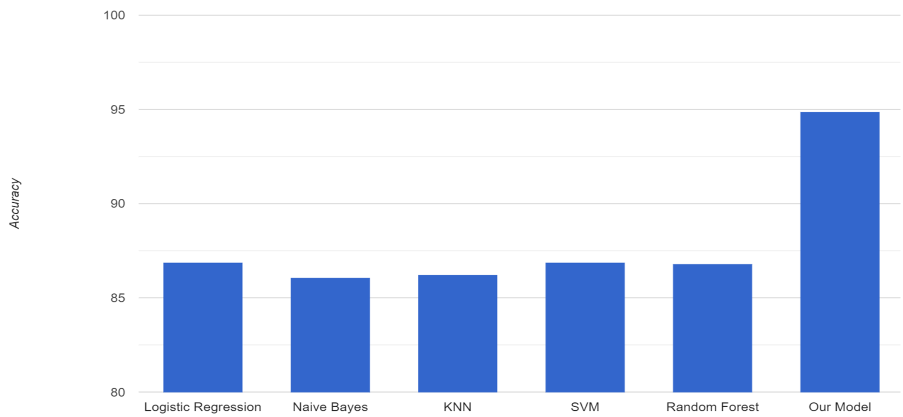

Fig 3. Comparison Chart with Traditional Models.

We have referred and analyzed from the paper "Spam review detection using deep learning"[9]. That uses CNN, LSTM and MLP for Spam review detection and its results are shown Fig. 4. Our model comparatively provides better Accuracy than their results. Our model was able to generalize and get good results on all kinds of validation data and testing data.

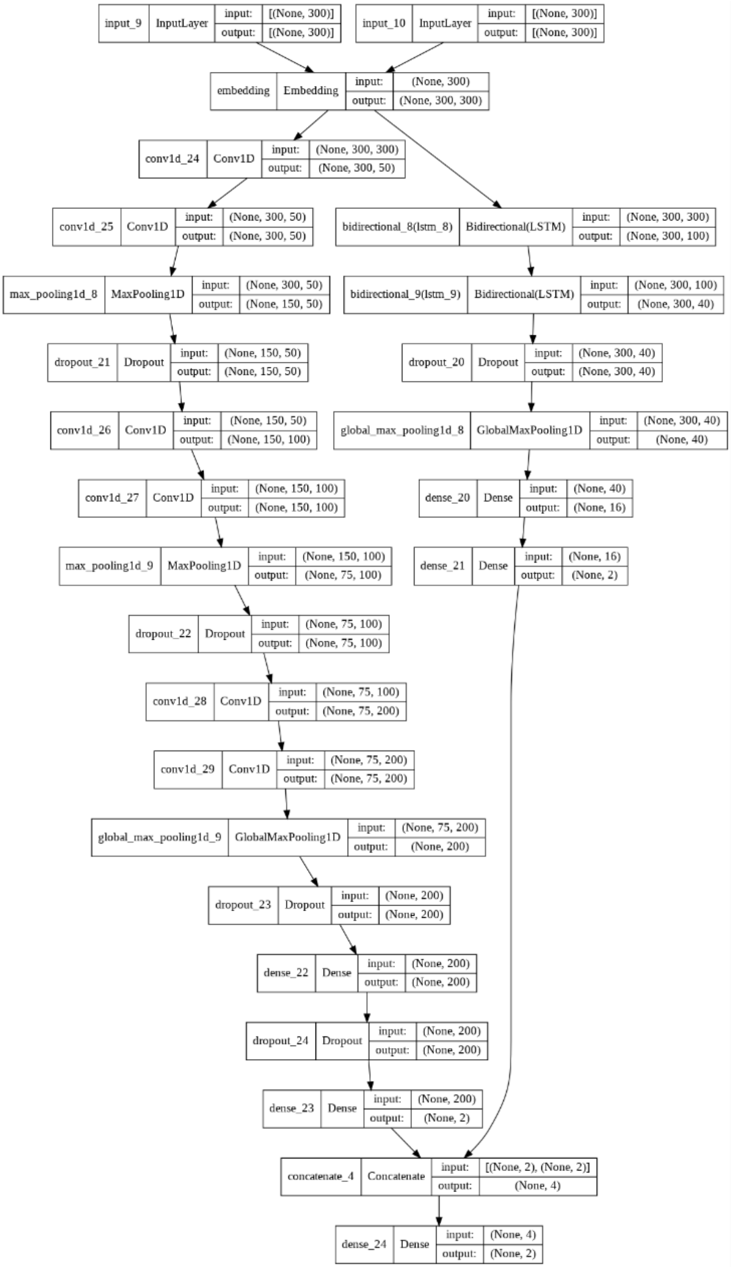

Fig 2. Model Architecture.

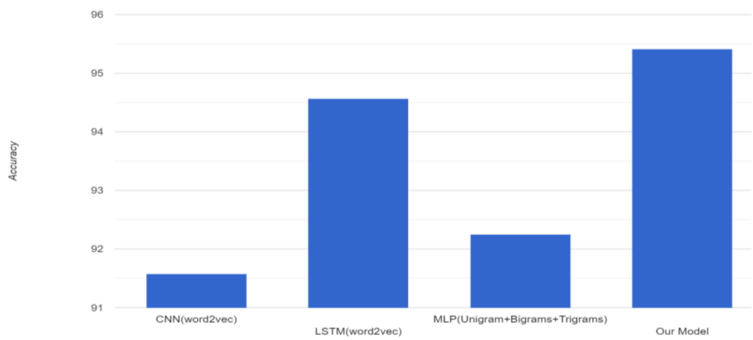

Fig. 4. Comparison Chart with Related works.

Our model scored around 95.41% accuracy on Validation data and 94.88% accuracy on Test data. The Fig 3 and 4. Shows the comparisons between related works and our proposed work. It can be observed that our model outperforms the existing models.

	precision	recall	f1-score	support
Fake	0.94	0.96	0.95	8799
True	0.96	0.94	0.95	8774
micro avg	0.95	0.95	0.95	17573
macro avg	0.95	0.95	0.95	17573
weighted avg	0.95	0.95	0.95	17573
samples avg	0.95	0.95	0.95	17573

Fig 5. Score Metrics.

For further analysis the standard measures like recall, f1-scores and precision are also measured. The model results are tabulated in Fig. 5. It is inferred from the table that the proposed model produces good results in all the evaluation metrics.

7. Discussion

With the advent of internet and evolution of technology ecommerce-based applications gained popularity in recent days. With more online purchases the product reviews plays an important role in the purchase of a product. Therefore, identifying fake reviews are of utmost importance and this paper proposes a deep learning model for fake review detection. We are able to achieve better accuracy of 94.88% which is 7.5% increase in accuracy to the existing work in Yelp dataset. The experimental results confirm that the proposed deep learning-based model produced exceptional results compared to the existing techniques. Still, there are a lot opportunities for the improvement to our work in the future. The number of reviews from Yelp Dataset can be increased and Attention Based Transformers Model can be developed for faster results and better accuracy.

References

[1] Jia, S., Zhang, X., Wang, X., & Liu, Y. (2018, May). Fake reviews detection based on LDA. In 2018 4th International Conference on Information Management (ICIM) (pp. 280-283). IEEE.

[2] Sanjay, K. S., & Danti, A. (2019, March). Detection of fake opinions on online products using Decision Tree and Information Gain. In 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC) (pp. 372-375). IEEE.

[3] Rodrigues, J. C., Rodrigues, J. T., Gonsalves, V. L. K., Naik, A. U., Shetgaonkar, P., & Aswale, S. (2020, February). Machine & deep learning techniques for detection of fake reviews: A survey. In 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE) (pp. 1-8). IEEE

[4] Elmogy, A. M., Tariq, U., Mohammed, A., & Ibrahim, A. (2021). Fake Reviews Detection using Supervised Machine Learning. Int. J. Adv. Comput. Sci. Appl, 12.Nidhi and Patel, 4th International Conference on Computing Communication and Automation (ICCCA) 978–979 (2018).

[5] Rajamohana, S. P., Umamaheswari, K., Dharani, M., & Vedackshya, R. (2017, March). A survey on online review spam detection techniques. In 2017 international conference on Innovations in green energy and healthcare technologies (IGEHT) (pp. 1-5). IEEE.

[6] Shuqin, Y., & Jing, F. (2019, August). Fake reviews detection based on text feature and behavior feature. In 2019 IEEE 21st International Conference on High Performance Computing and Communications; IEEE 17th International Conference on Smart City; IEEE 5th International Conference on Data Science and Systems (HPCC/SmartCity/DSS) (pp. 2007-2012). IEEE.

[7] Hassan, R., & Islam, M. R. (2021, February). Impact of Sentiment Analysis in Fake Online Review Detection. In 2021 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD) (pp. 21-24). IEEE

[8] Patel, N. A., & Patel, R. (2018, December). A survey on fake review detection using machine learning techniques. In 2018 4th International Conference on Computing Communication and Automation (ICCCA) (pp. 1-6). IEEE.

[9] Shahariar, G. M., Biswas, S., Omar, F., Shah, F. M., & Hassan, S. B. (2019, October). Spam review detection using deep learning. In 2019 IEEE 10th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON) (pp. 0027-0033). IEEE.

[10] Liu, M., Shang, Y., Yue, Q., & Zhou, J. (2020). Detecting Fake Reviews Using Multidimensional Representations With Fine-Grained Aspects Plan. IEEE Access, 9, 3765-3773.

[11] Wahyuni, E. D., & Djunaidy, A. (2016). Fake review detection from a product review using modified method of iterative computation framework. In MATEC web of conferences (Vol. 58, p. 03003). EDP Sciences.

[12] Patel, D., Kapoor, A., & Sonawane, S. (2018). Fake review detection using opinion mining. International Research Journal of Engineering and Technology (IRJET), 5.

[13] Hajek, P., Barushka, A., & Munk, M. (2020). Fake consumer review detection using deep neural networks integrating word embeddings and emotion mining. Neural Computing and Applications, 32(23), 17259-17274.

[14] Wang, J., Kan, H., Meng, F., Mu, Q., Shi, G., & Xiao, X. (2020). Fake review detection based on multiple feature fusion and rolling collaborative training. IEEE Access, 8, 182625-182639.

[15] Abhinandan, V., Aishwarya, C. A., & Sultana, A. (2020). Fake Review Detection Using Machine Learning Techniques. International Journal of Fog Computing (IJFC), 3(2), 46-54.