2019 International Conference on Advance in ICT for Emerging Region (ICTer) : 1 - 6

Opinion Spam Detection in Online Reviews Using Neural Networks

K. Archchitha[#1], E.Y.A. Charles[#2],
[#]Department of Computer Science, University of Jaffna
Jaffna, Sri Lanka
[1]archchithak@gmail.com
[2]charles.ey@univ.jfn.ac.lk

Abstract— **People of the Internet era usually rely on online reviews to make decisions about an online purchase, a hotel booking or a car rental and much more, since people believe that making decisions based on other's opinions lead to making the right choice. As writing fake reviews come with monetary gain, the opinion spam activities have increased dramatically on online review websites. Thus opinion spam in reviews has become a big challenge to people to make purchase decisions as well as damages the reputation of review websites. Hence the deceptive opinion spam detection is an essential task in the field of natural language processing. Most of the existing research on opinion spam detection uses the traditional bag-of-words model to represent the review text features and apply standard machine learning models such as Support-Vector Machines or Naïve Bayes as classifiers. There is only a few state of the art methods, which have utilized neural network based methods for spam detection. With the recent advancements in deep learning, there are successful applications of Convolutional Neural Networks (CNN) for natural language processing problems which have achieved improved performance. In this study, a CNN model is developed to detect opinion spam using the features extracted from the pre-trained GloVe: Global Vectors for Word Representation model. Moreover, some word and character level features used in existing research work, are extracted from the text and concatenated with a feature set extracted by the convolutional layers of the model to improve the performance. The proposed model found to outperform the state-of-the-art method and the inclusion of additional features improved the performance further.**

Keywords— **Opinion spam detection, Natural language processing, Deep learning, Convolutional neural network.**

I. INTRODUCTION

Online reviews play a very important role in decisions making of people and influence their daily purchase decisions. People often read reviews of products and services before making a decision on whether to buy/book it or not or where to buy/book it. Fake or fraudulent or deceptive reviews written by malicious users are the opinions that could mislead the customers. The fake review writers are the people who may not have bought or used a particular product/service but hired by competitors for writing reviews to deceive the customers [1]. Positive reviews of target entities attract more customers and increase sales while negative reviews damage the reputation of the particular entity and decrease its reputation. Reading the opinions on Web has become very common for people to fulfil various decision-making issues. Many individuals and organizations are increasingly using online reviews for their decision making, hence fake reviews have become a big challenge to them. Fake reviews also damage the reputation of online review websites. Since most of the recommendation systems are rely on the correctness of user reviews, these opinion spams decrease the trustworthiness of the recommendation systems as well. Hence, opinion spam detection is a very essential research area which can also be considered as a very active research field of natural language processing. This research work aims to identify whether a review is fake or not by using an opinion spam detection model and to analyze the effect of various text features on opinion spam identification.

Previous work on spam detection has utilized a variety of features extracted from opinion spam document corpora. Some of the previous work on the opinion spam detection utilized the n-gram based text categorization features [2, 3]. Some prior studies considered review writers' behavioural patterns [4, 5, 6, 7]. Moreover, review content-related features such as lexical features, semantic inconsistency, review length, content and style similarity are also used for this purpose in some studies [1, 8]. Most of the prior research applied a traditional bag of word approach to representing the review text features and used classifiers such as linear regression, Support Vector Machines (SVM) and Naïve Bayes. In the bag-of-word representation, only the meaning of the words is considered and the association between them are not adequately represented. Word Embedding is a way to represent words as dense vectors that also inherently capture the context of a word and semantic & syntactic similarity relations with other words. Therefore, word embedding is used as a high-level feature representation method in this study to achieve better accuracy.

Deep learning models have achieved remarkable results in computer vision and speech recognition. Within natural language processing, much of the work with deep learning methods have involved learning word vector representations through neural network models and performing composition over the learnt word vectors for classification [9, 10, 11]. Deep learning models have better self-adaptability, which make the models scalable and also suitable for online learning scenarios where the model needs to get updated continuously with new data. Convolutional Neural Networks (CNN) utilize layers with convolving filters which are applied to extract local features. CNN models, originally invented for Computer Vision applications, have subsequently been shown to be effective for Natural Language Processing (NLP). It has achieved excellent results in semantic parsing, search query retrieval, sentence modelling, and other traditional NLP tasks [9]. Compared with traditional n-gram methods, CNN is also found to be efficient in terms of feature representation. In CNN feature extraction is performed by

978-1-7281-5156-4/19/$31.00 ©2019 IEEE

Convolution layers. Convolution is a set of mathematical operation to extract features, can be efficiently implemented on a hardware level through GPUs using open source libraries such as TensorFlow. Since CNN architectures have obtained state-of-the-art results on many text classification studies, this research work proposes a classification model based on the CNN model introduced by Yoon Kim [9]. This study proposes a simple CNN with three parallel convolution layers with different filter sizes, which is trained using word vectors obtained from Global Vectors for Word Representation (GloVe), a pre-trained word vector model.

Online reviews are quite short with various types and contents. To effectively identify misleading opinions in the reviews, there is a need to comprehensively study the characteristics of misleading opinions and examine relevant features in addition to textual semantics [12]. The proposed approach combines features extracted by deep learning and traditional text and behavioural features using a multilayer neural network which acts as a meta-classifier. The proposed model was implemented using TensorFlow and demonstrated improved accuracy. The proposed model can be applied to other text categorization tasks such as sentiment analysis and auto-tagging of customer queries.

II. LITERATURE REVIEW

In recent years opinion spam detection is an active field of research, due to the dramatic increase in on-line purchase/booking and active usage of user opinions for decision making. Researchers have been proposed many techniques and approaches in this field. This section presents a brief review of the existing work.

A. Opinion spam detection

Jindal and Liu [1] have introduced the idea of opinion spam detection and have claimed that opinion spam requires different detection techniques since it is quite different from Web spam and email spam. They proposed a behavioural feature-based method for opinion spam detection. Their experiments were conducted on a large Amazon review dataset with the details of Product ID, Reviewer ID, Rating, Date, Review, Title, Review Body, Number of Helpful Feedbacks and Number of Feedbacks. In this work, the features were categorized into three types of features, namely, review centric features, reviewer centric features and product centric features. The proposed model considered spam detection as a classification problem with two classes, spam and non-spam and have built supervised learning models to detect individual fake reviews.

Ott et al. [2] proposed a fake review detection method mainly considering the review text information. The work was based on all 5-star truthful reviews for 20 most famous hotels in the Chicago area from Tripadvisor and deceptive opinions gathered for the same hotels using Amazon Mechanical Turk (AMT). Fake review detection was carried out using three strategies - genre identification, detection of psycholinguistic deception, and text categorization. Proposed method focused on the part-of-speech (POS) distribution of the reviews, particularly the frequency of POS tags. For prediction, classification methods such as Support Vector Machine and Naive Bayes Classifiers were used. Collected reviews were first evaluated by human judges and the results were compared with the results obtained by the automated classifiers. It was found that automated classifiers outperformed humans for each metric. Further, this work derived a conclusion that the standard n-gram–based text categorization performed better for fake review detection, but a combination approach using psycholinguistically motivated features and n-gram features can perform slightly better.

Vlad Sandulescu and Martin Ester [3] employed the Latent Dirichlet Allocation (LDA) model using bag-of-words and bag-of-opinions to detect singleton or one-time review spammers. LDA model is known to be operating on semantic similarity of text. Features used were topics extracted from the review texts with parts-of-speech (POS) patterns.

Rayana and Akoglu [4] proposed a model which combines features from review texts, time-stamps, user behavioural information and the review network under a unified framework to spot suspicious users and reviews, as well as targeted products of spam.

Mukherjee et al. [5] proposed a model for spotting fake reviewer groups in consumer reviews that uses a frequent itemset mining method to find a set of candidate groups. They used several behavioural models derived from the collusion phenomenon among fake reviewers. In addition, they used relation models based on the relationships among groups, individual reviewers, and the products they reviewed. And also, they built a labelled dataset of fake reviewer groups. Their method outperformed multiple strong baselines including the state-of-the-art supervised classification.

Song Feng et al. [8] have investigated syntactic stylometry for deception detection from four different datasets. They applied the features driven from Context Free Grammar (CFG) parse trees and used SVR classifiers.

B. Artificial Neural Networks for opinion spam detection

Apart from standard machine learning models, some researchers have used Artificial Neural Network (ANN) models for fake review detection and obtained better performances and some interesting results.

Manqing Dong et al. [7] introduced an end-to-end trainable unified model to leverage the appealing properties from Autoencoder neural networks and random forest. They implemented a stochastic decision tree model to guide the global parameter learning process on large Amazon review dataset.

Kunal Goswami et al. [6] used ANN to evaluate the reviewer social interaction feature set of Yelp and found that the features, namely, Maximum number of reviews, Number of firsts, Average rating distribution, Review count, Tips, Photo Count, Friend Count, Cool Votes, Followers, Useful Votes, Burstiness, Funny Votes, Negative review ratio, Positive review ratio, ERD, ETG, Compliments and Bookmarks are the most important features in the priority order for opinion spam detection. Wang et al. [13] designed a semi-supervised recursive autoencoder for detecting review spam in Weibo, a Chinese alternative to Twitter.

C. *Convolution neural networks for text classification*

Yoon Kim [9] described a series of experiments with convolutional neural networks (CNN) trained on top of pre-trained word vectors (word2vec) for sentence-level classification tasks. They have shown that a simple CNN with little hyperparameter tuning and static vectors can achieve excellent results on multiple benchmarks. The proposed CNN models have shown improved performance comparing to the state-of-the-art models on 4 out of 7 tasks. Yoon Kim also experimented with two different static and dynamic word embedding channels, where he adjusted one channel during training and other was not adjusted.

A similar, but somewhat more complex, architecture was previously proposed in [14]. They describe a convolutional architecture which uses Dynamic k-Max Pooling, a global pooling operation over linear sequences. The network handles multiple lengths of sentences as input and explicitly capturing short and long-range relations.

In machine learning, an entity is often represented using a set of numerical features. For machine learning related to text processing, each word is a feature and is often represented by a single value or set of values such as term frequency (TF) or inverse document frequency (IDF). This type of representation is considered inadequate and currently distributed word representation models such as Word2Vec or GloVe are increasingly used for machine learning for text processing. Here, the basic idea is to represent each word with a dense vector which inherently captures the context of a word, semantic and syntax similarity and relationship with other words. Global Vectors for Word Representation (GloVe) is a publically available data which was obtained by training on aggregated global word-word co-occurrence statistics from several corpora. However, the GloVe is different from other representation methods which gives a vector space with meaningful sub-structures [20].

Rie Johnson and Tong Zhang [15] introduced a different model for text categorization that trained a CNN from scratch instead of using pre-trained word vectors like Word2Vec or GloVe. They applied convolutions directly to one-hot vectors.

Ye Zhang et al. [11] performed an empirical study on the effect of varying hyperparameters in CNN architectures. They focused on one layer CNNs and investigated the impact on performance and variance over multiple runs.

D. *Features for opinion spam detection*

In the past, researchers have used a large number of features for opinion spam detection. These features can be categorized into four types – Review content-based features, Reviewers' behavioural features, Reviewed product-related features and Relational features. The review content features include lexical features, semantic inconsistency, review length, content and style similarity. Behavioural features are derived from the available public data and web sites private/internal data. These features include reviewer id, time of posting, frequency of posting, first reviewers of products, IP and MAC addresses, time taken to post a review and physical location of the reviewer. Product related features include description, sales volume, and sales rank. Complex relationships among reviewers, reviews and entities (e.g., products and stores) falls under the Relational features. When comparing the results obtained from the existing experiments with these four types of features, it can be concluded that the review content-based features are most important and useful for opinion spam detection models.

E. *Classification techniques*

Automated opinion spam detection is a machine learning application and previous work in this area has reported the applicability and outcome of several classifiers for this task. Much of the existing research work used logistic regression (LR), linear support vector machines (SVM) and naïve Bayes (NB) classifier. Jindal and Liu [1] compared these three classifiers and found that LR performs best on their data. Ott et al. [2] and Mukherjee et al. [5] on the other hand, found that SVM with a linear kernel yields the best performance. Li et al. [21] used a semi-supervised co-learning algorithm besides a traditional naïve Bayes classifier to make use of the large pool of unlabeled data and found that this model outperforms NB by a relatively large margin. Akoglu et al. [16] in particular represent the opinion spam detection as a bipartite graph where products and reviewers are represented as nodes and reviews as graph edges. There are also some reported works using semi-supervised approaches for opinion spam detection. Some researchers have considered text analysis as a task of learning the written patterns. In these research works, probabilistic methods such as unsupervised Bayesian approach [5] and Hidden Markov models were used for the above mentioned task.

There are several studies on opinion spam detection using deep learning methods. Zhao et al. [12] proposed a word-order preserving CNN model. Yafeng et al. [17] proposed a CNN model which extract information about sentence representation combined with discourse information. Even though these works have claimed that their proposed models outperform state of the art Tf-Idf based models, the reported accuracies were observed to be low.

III. METHODOLOGY

Based on the literature review, it is found that state-of the-art methods give much consideration for individual words. As an alternative, this study utilizes Word Embedding which represents words as dense vectors. Since CNN architectures [9, 18, 19] have obtained the state-of-the-art results on many text classification datasets, this study proposes a model based on a recently proposed CNN model in [9]. Following is the stages of the opinion spam detection model proposed in this study:

- Data preprocessing

  As the first step, the reviews in the data set are cleaned and punctuations removed. Then each review text is transformed into a sequence of tokens and the text sequences are zero-padded to make all the sequences to have fixed length M.

- Feature Representation

  Tokens are mapped to their respective embedding using the GloVe pre-trained word embedding model. Each taken is represented with a vector of N dimension. This approach differs from the bag-of-word approach which represents semantic information in fewer dimensions.

- Model construction & Training

  For the said purpose, CNN model with three parallel convolution layers with different filter sizes is created. The model is trained using the data set of review text where each review text is represented by a vector of M x N dimension. Here M is the length of review text in words and N is the dimension of a word vector from GloVe.

- Performance improvement

  Additional features from review text are extracted and concatenated with the flattened output of the convolution layers in the CNN model. In order to improve the performance, this combined feature set is used to train the model for classifying review text.

A. Data preprocessing

In text classification problems, basic preprocessing operations are needed to reduce the computational expense. Here, text cleaning and punctuation removal techniques are applied as basic preprocessing operations. Tokenization of sentences is one of the essential parts in natural language processing that simply divides a sentence into a list of words. In this research study, a collection of labelled review text is used to train a machine learning model. The review text documents are transformed into sequences of tokens. Large review text sequences are truncated to a maximum length of 1000 words. Review texts with fewer words are zero padded to have an equal length as other review texts.

B. Feature representation

To use text documents for training the proposed CNN machine learning model each of the preprocessed reviews is converted into numerical features. The pre-trained GloVe model provides a dense vector of dimensions 50, 100 and 300 for each English word in a very large corpus. Each word in each of the tokenized text documents is mapped to the respective dense vector of 300-dimensional pre-trained GloVe word embedding.

C. CNN Model

The CNN model proposed in this study consists of three parallel convolution layers with different filter sizes. After each convolution layers, a standard max-pooling operation is performed on the latent space. The max-pooled output from the three parallel streams is concatenated and fed to another two convolution layers. Each convolution layer consists of a convolution, a max-pooling and a dropout operation with a specific dropout value (0.6). The output from the convolution layers is flattened and fed to a fully connected layer with ReLU activation function. The output of this layer is fed to a fully connected layer with a sigmoid activation function. This is the final layer which performs the classification. Figure 1 shows a schematic diagram of the proposed CNN model.

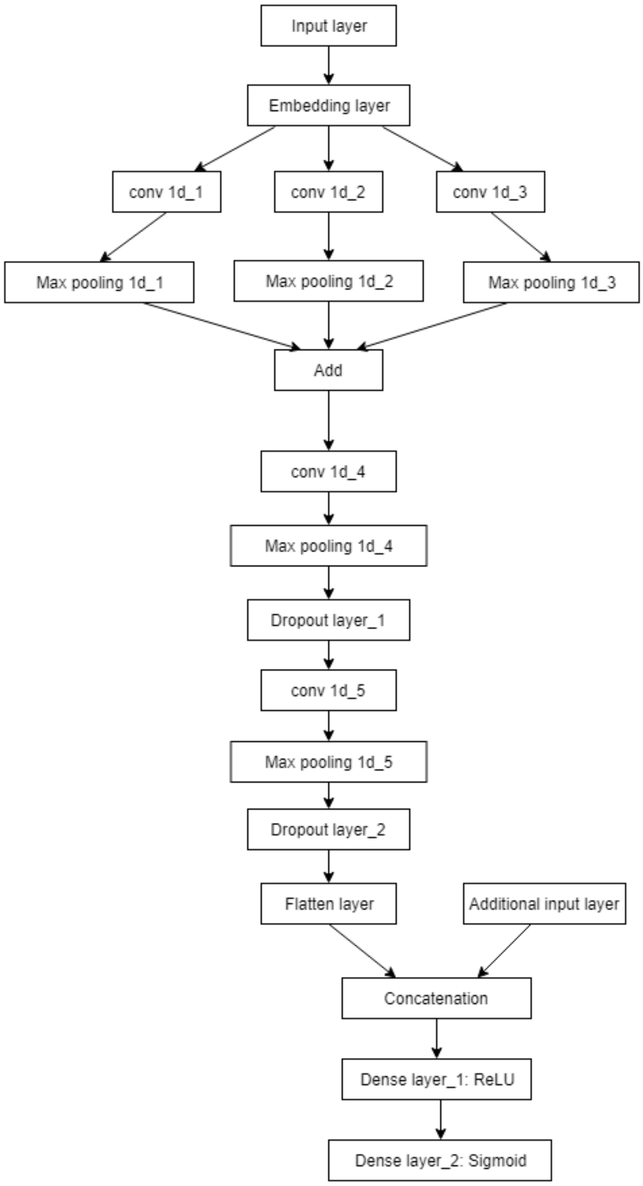

Fig. 1 Proposed CNN model for fake review detection

D. Performance Improvement

The study also tried to verify the effect of incorporating the traditional text-based features in addition to the word embedding features. Features tested are number of words, number of characters, number of stop words, number of special characters, number of upper cases and some part of speech tags frequencies. The selected features are extracted from the data set separately and concatenated with the features obtained from the flatten layer. Finally, the concatenated features are fed to the fully connected layer with a sigmoid activation function to build a classifier.

IV. EXPERIMENTAL SETUP & TEST RESULT

A. Dataset

To evaluate the performance of the proposed model, it was implemented, trained and tested. For the training and testing, the Deceptive Opinion Spam Corpus created by Ott et al. [2] is used. This corpus consists of truthful and deceptive hotel reviews of 20 Chicago hotels with 400 truthful positive reviews from TripAdvisor, 400 deceptive positive reviews from Amazon Mechanical Turk, 400 truthful negative reviews from Expedia, Hotels.com, Orbitz, Priceline, TripAdvisor and Yelp and 400 deceptive negative reviews from Amazon Mechanical Turk. All the reviews in the corpus fall into two categories, truthful or deceptive and are labelled. Table I list the details of the corpus [22]. Out of the full corpus of 1600 reviews, 75% was used for training and 25% was used for testing. Even though the dataset consists of positive and negative feedback, the study concentrated on identifying whether a review is truthful or deceptive.

TABLE I
DATASET DETAILS

No. of reviews	**Type**	**Source**
400	Truthful positive	TripAdvisor
400	Deceptive positive	Amazon Mechanical Turk
400	Truthful negative	Expedia, Hotels.com, Orbitz,Priceline, TripAdvisor and Yelp
400	Deceptive negative	Amazon Mechanical Turk

B. Experimental Setup

The proposed CNN model was implemented using Keras open source library. Keras is a deep learning and neural networks API which is providing high-level building blocks for developing deep-learning models. It is a model-level open source library and is capable of running on top of Tensorflow. The implementation also used Adam, one of the most common optimization algorithm currently used. Adam is an extension to stochastic gradient descent that has recently seen broader adoption for deep learning applications in computer vision and natural language processing with good performance improvements [23]. Further, the binary cross-entropy which is used for binary classification problems is used as the objective function which is often referred to as a cost function or a loss function to minimize the error. To evaluate the accuracy of the proposed models, test score and confusion matrices are used.

C. Test Results

A number of experiments were conducted to evaluate performance of the proposed CNN model to classify opinion reviews. To compare the results obtained from various approaches and experiments, two baseline models were built as the initial step. One base line model was built using count vectors as features, since most of the existing research work used this feature. Here the count vector is a simple numerical representation of the review text dataset. The count vectors are made up with the frequency of each word in each sentence. The other baseline model was built using TF-IDF features. TF-IDF features represents the relative importance of a term in a review document and in the entire corpus. These initial models were built using input as word level n-gram values for n=1, 2 and 3. The best accuracy was selected from the results obtained from these experiments.

As the next stage, the proposed CNN model was trained and tested with the word vectors as features. After several round of tuning the proposed CNN model, it was able to predict the type of the opinion review with an accuracy of 86.25% when word embedding features were used. The model correctly identified 183 deceptive reviews and 165 truthful reviews while it failed to detect 10 deceptive reviews and 42 truthful reviews. Figure 2 shows the change in training and validation accuracy as well as the training and validation loss during the training of the proposed model. Table II shows the confusion matrix obtained for the model.

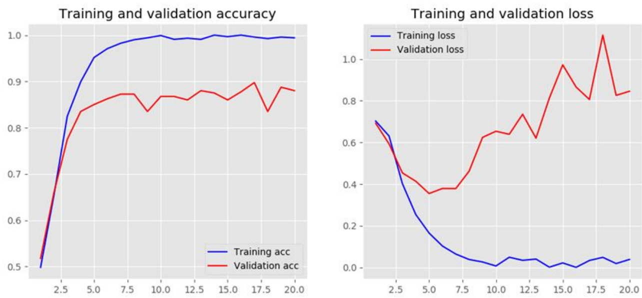

Fig. 2 Change of accuracies and loss with training for the CNN model

TABLE II
CONFUSION MATRIX FOR THE CNN MODEL TRAINED WITH WORD EMBEDDING FEATURES

		Predicted label	
		Truthful	Deceptive
Actual label	Truthful	165	42
	Deceptive	10	183

To analyse the effect of various types of features, additional CNN models were built with a combination of word embedding features and text-based features. A model was built using word embedding and the following text-based features: Number of words, number of characters, number of stop words, number of special characters, number of upper cases and some part of speech tag's frequencies. Another CNN model was built using word embedding features and TF-IDF features. One more CNN model was built using word embedding features count vector features. Table III lists the accuracies obtained for various trials described above.

It can be observed from the experimental results, that the concatenation of the additional are contributed to improving the performance of the proposed model to some extent, even though the increase is low. Text-based features improved the performance to 88.25%, TF-IDF features improved the performance to 88.75% and the concatenation of count vectors to 88.50%.

TABLE III
EXPERIMENTAL RESULT

Features	Training accuracy	Testing accuracy
Count vectors as features	99.50%	81.25%
TF-IDF as features	99.25%	80.50%
Word embedding as features	100.00%	86.25 %
Word embedding + additional features	99.58%	88.25%
Word embedding + TF-IDF features	99.92%	88.75%
Word embedding + Count vectors features	99.83%	88.50%

V. DISCUSSION

The proposed model learns was observed to be distinguish truthful and deceptive reviews using latent features of reviews in comparison to observable text-based features. Compared with traditional learning models, CNNs are fast and also efficient in terms of representation of the domain. For text related tasks, the performance of traditional approaches degrades with the increase in the coverage of the task domain due to the huge increase in the vocabulary. Unlike this, convolutional filters learn good representations automatically, without needing to represent the whole vocabulary. Hence the proposed CNN model with word embedding makes the model scalable, which is a needed property for text based learning tasks. The proposed model was able to perform with accuracy above 88%, where the number of reviews used is 1600. But in real world an online review portal contains large number of reviews. It is well known that CNNs work best with large training sets where they are able to find good generalizations where a simple model like logistic regression will not be able. By performing experiments with a various combinations of type of features showed some interesting results. By concatenating the text-based features the performance of the proposed CNN model was able to be increased by some small extent. It needs to be analysed further about the proportion of computational resource increase compared to the increase in performance.

VI. CONCLUSIONS

In this study, a CNN is proposed to distinguish truthful opinion reviews from opinion spams using. The proposed model was trained with the opinion text which were represented as word vectors using the pre-trained GloVe word embedding model. This approach differs from previous approaches in two ways. Most of the previous approaches used the traditional bag-of-words model to represent the review text features and relied on classifiers such as Support-Vector Machines and Naïve Bayes. The proposed CNN model was tested on Deceptive opinion spam corpus and found to be outperform traditional approaches. The experiments have shown that in order to effectively identify misleading opinions in the reviews, it is necessary to examine new features in addition to textual semantics. The Proposed approach is able to combine both deep learning and traditional feature-based approaches using a multilayer neural network which acts as a meta-classifier. The proposed model demonstrated its capability by improving the accuracy further when combined with the traditional text based features. The performance of the proposed model can be further improved by utilizing a larger data set, using additional features such as behavioral features and also by fine tuning the hyper-parameters of the CNN model. The classification approach is similar to many text based machine learning tasks. Hence, this model also can be applied to tasks such as sentiment analysis, auto tagging of customer queries, and categorization of text into defined topics.

REFERENCES

[1] Nitin Jindal and Bing Liu, “Opinion spam and analysis”, 2008.
[2] Myle Ott, Yejin Choi, Claire Cardie and Jeffrey T. Hancock, “Finding Deceptive Opinion Spam by Any Stretch of the Imagination”, 2011
[3] Vlad Sandulescu and Martin Ester, “Detecting Singleton Review Spammers Using Semantic Similarity”, 2015
[4] Shebuti Rayana and Leman Akoglu, “Collective Opinion Spam Detection: Bridging Review Networks and Metadata”, 2015
[5] Arjun Mukherjee, Bing Liu and Natalie Glance, “Spotting Fake Reviewer Groups in Consumer Reviews”, 2013
[6] Kunal Goswami, Younghee Park and Chungsik Song, “Impact of reviewer social interaction on online consumer review fraud detection”, 2017
[7] Manqing Donga, Lina Yaoa, Xianzhi Wanga, Boualem Benatallaha, Chaoran Huanga and Xiaodong Ninga, “Opinion Fraud Detection via Neural Autoencoder Decision Forest”, 2018
[8] Song Feng, Ritwik Banerjee and Yejin Choi, “Syntactic Stylometry for Deception Detection”, 2012
[9] Yoon Kim, “Convolutional Neural Networks for Sentence Classification”, Empirical Methods in Natural Language Processing (EMNLP),2014
[10] William Yang Wang, “Liar,LiarPantsonFire”: New Benchmark Dataset for Fake News Detection”, 2017
[11] Ye Zhang, Byron C. Wallace, “A Sensitivity Analysis of (and Practitioners’ Guide to) Convolutional Neural Networks for Sentence Classification”, 2016
[12] Siyuan Zhao and Zhiwei Xu and Limin Liu and Mengjie Guo, “Towards Accurate Deceptive Opinion Spam Detection based on Word Order-preserving CNN”, 2017.
[13] Wang, B., Huang, J., Zheng, H. and Wu, H., “Semi-supervised recursive autoencoders for social review spam detection”, 2016
[14] Nal Kalchbrenner, Edward Grefenstette, Phil Blunsom, “A Convolutional Neural Network for Modelling Sentences”, 2014.
[15] Rie Johnson, Tong Zhang, “Effective Use of Word Order for Text Categorization with Convolutional Neural Networks”, 2014
[16] Leman Akoglu and Hanghang Tong and Danai Koutra, Graph based anomaly detection and description: a survey, Data Mining and Knowledge Discovery, Vol. 29, pages 626-688, 2014.
[17] Ren, Yafeng & Ji, Donghong, “Neural networks for deceptive opinion spam detection: An empirical study”, 2017.
[18] Xiang Zhang, Junbo Zhao, Yann LeCun, “Character-level Convolutional Networks for Text Classification”, 2016.
[19] Ronan Collobert, Jason Weston, “A Unified Architecture for Natural Language Processing: Deep Neural Networks with Multitask Learning”, 2008.
[20] Jeffrey Pennington, Richard Socher, Christopher D. Manning, “GloVe: Global Vectors for Word Representation”, in Empirical Methods in Natural Language Processing (EMNLP), pages:1532-1543, 2014.
[21] Luyang Li, Bing Qin, Wenjing Ren and Ting Liue, “Document representation and feature combination for deceptive spam review detection”, 2017.
[22] Deceptive Opinion Spam Corpus - https://myleott.com/op-spam.html
[23] Diederik P. Kingma and Jimmy Ba, “Adam: A Method for Stochastic Optimization”, CoRR, 2015.