# Fake News Detection in Social Networks Using Machine Learning and Deep Learning: Performance Evaluation

Wenlin Han* and Varshil Mehta
*Department of Computer Science*
*California State University, Fullerton, USA*
whan@fullerton.edu; mehtavarshil@csu.fullerton.edu

***Abstract*—The problems related to fake news are growing rapidly which results in misleading views on some information. Social media networks are one of the fastest medium to spread information by creating a huge impact on manipulating information by influencing readers in positive and negative aspects. This paper aims at evaluating and comparing different approaches that are used to mitigate this issue including some traditional machine learning approaches, such as Naive Bayes, and the popular deep learning approaches, such as hybrid CNN and RNN. The comparison is not only within traditional methods or within deep learning methods, but also across traditional and non-traditional methods. This paper lays a foundation for selecting a machine learning or deep learning method for problem solving regarding the balance between accuracy and lightweightness.**

***Index Terms*—Fake News Detection, Deep Learning, Machine Learning, Natural Language Processing, Social Media**

## I. Introduction

In 2016 following the election process, there were a lot of articles having factually incorrect data and misleading information for making money through page views. Spreading fake news results in distrust which creates biased reviews leading biased perception. Fake news posses' negative effects on people's minds making them read the news that is less credible. Such misleading news information is spread on that basis of making money on the number of views for that page or posing biased opinion that tries to create hoax among people's mind that leverages into a change in their decision-making like in case of political election [6].

As compared to real news, fake news spreads more quickly by making it more appealing on viewer's end. It gets difficult to distinguish between fake news and real news. The human tendency is to believe that the knowledge of their facts is precise that leads to bias views on information making it recipients for "yellow journalist". This reveals the information that would be considered by any person based on his/her knowledge which leverages to their perspective and their biased views. With the help of social media sites as a tool to spread rumors and companies like Facebook, Twitter have faced scrutiny in past times. Fake flag news feature was adopted by Facebook [6]. Social media platform gives access to news and information online to share and comment which is less expensive as compared to traditional news media. For example, in 2016 about 62 percent of U. S. adults get news on social media compared to only 49 percent of adults reported watching the news on social media. It is important to verify the origin of the information and to filter out the fake information to the authentic ones [5].

Based on the input sources fake news method are classified into two categories news and social context models. For news content the contents of news can be differentiated between two categories which are Linguistic based (including texts, heading, etc.) and Visual based (including video and image-based). By comparing different approaches based on categories such as news content model which focuses on the body-text, title and analyzing the content of the news on how interaction is done for some preferred news on social media. The method for social context focuses on social features and signals which is the interaction of users with given news on social media referred to as social-based methods [1]. The content-based method is traditional approach focusing on conventional news media which has no social information is available.

The main contributions of this paper include:

- summarizing fake news types and features;
- evaluating and summarizing fake news detection performance for traditional machine learning approaches;
- evaluating and summarizing fake news detection performance for the latest deep learning approaches;
- comparing performance between traditional machine learning and the latest deep learning approaches;
- laying a foundation for selecting a machine learning or deep learning method for problem solving regarding the balance between accuracy and lightweightness.

The rest of the paper is organized as follows: various types of fake news are studied and summarized in Section II. In Section III, we present performance evaluation of some traditional machine learning detectors. In Section IV, we introduce fake news detection performance regarding TF-IDF and PCFG. The performance evaluation of popular deep learning methods are presented in Section V. We compare the performance of the traditional machine learning with deep learning methods in Section VI. Finally, we conclude the paper in Section VII.

## II. News and Fake News

The contents of any news are provided by carrier platforms to its end users - a jump of 5 percent which is around two-thirds adults in the U.S. get their news from social media in 2017 as compared to 2016. Here four major different formats user consume information shown in Table I.

978-1-7281-2977-8/19/$31.00 ©2019 IEEE
DOI 10.1109/ICII.2019.00070

TABLE I
DIFFERENT FORMAT HUMANS CONSUME INFORMATION

| | |
|---|---|
| Multimedia | The combination of different contents such as audio, video, images which can be stored, processed and transmitted digitally. |
| Text | The content is analyzed by test linguistics, focusing on the part of communication in text format. Other characteristics are also focused like grammar, tone. |
| Audio | Apart from being part as multimedia, it has a separate medium as a news source which includes broadcast network, radio, and podcast making its end users spectrum more. |
| Hyperlink | Linking on to the different aspects of the information depending upon the writer to gain readers trust. |

In a very general context, researchers of social science have studied fake news, and its types are categorized as shown in Table II.

The researchers have also studied the reliable and unreliable media sources. The results [6] are shown in Table III.

## III. TRADITIONAL MACHINE LEARNING METHODS

Fake news method detection mainly relies on feature extraction. Here, following approaches based on feature execution are categorized as shown in Table IV.

### A. Deception Modeling

Theoretical approaches for deceptive vs. truthful stories clustering process: Vector Space Modeling (VSM) and Rhetorical Structure Theory (RST) [2]. In RST, a hierarchical structure of each story by procedural analysis of capturing the logic of a story in the context of functional relations among different meaning of text units. Also, empirical research shows that to express certain idea writers tend to emphasize certain parts of the text. RST theory systematically identifies emphasized parts of text using rhetorical connections.

From RST resulted sets, VSM identifies relations of rhetorical structure. By representing in high dimensional space of news text and for various computer algorithm requires extracted text to be modeled in a suitable manner. In a set of the news report, the number of rhetorical relations is referred by the vector space of each dimension. Further analysis can be performed as this representation makes vector space very applicable and simple.

From the result of each text which is analyzed and converted to a set of rhetorical connection in a hierarchical tree by applying RST. The results of the rhetorical structure are identified using VSM.

### B. Clustering

To contrast and compare a large amount of data using clustering. Clustering package like gCLUTO (Graphical Clustering Toolkit) which differentiates news reports based on the fact of being similar on a chosen clustering algorithm. By using agglomerative clustering with a k-nearest approach for sorting/sorting a small number of clusters and a large number of data set. This approach clusters news report which is similarly based on their normalized frequency relations. This model measures the deceptive value of news based on the principle of coordinated distances. After computing non-deceptive and deceptive clusters centers, based on the deceptive values of the Euclidean distances from the center new incoming news were assessed. Though due to achieving lesser accuracy, this method is not a useful approach to the large data set. A similar story sets might not available to a recent fake news story which might not produce accurate results.

### C. Naive Bayes

By considering identical common properties between spam messages (i.e., irrelevant or messages that are unsolicited sent) and news articles which are fake such as

- Grammatical mistakes
- They are often construed using the emotional expression
- Manipulating opinion on some topics which affects the reader's opinion.
- From a syntactic point of view, have similarity on a limited set of words used as spam messages have similarity compared to different spam messages. By looking from a syntactic point of view, news articles which are fake and spam messages have the same similarity.
- Content on both fake news and spam messages are not true (in context of most spam messages and news which is fake often hold similarity)

As spam messages and fake news articles have a lot of properties which are common, so by using similar approaches to filter out fake news and spam messages.

Naive Bayes classifiers [3] are based on applying Bayes theorem which is assuming independence among the predictors. It shows independence between features present in the class, and if they are dependent on each other than or dependent on the existence of some other feature, then the probability is independently contributed from the properties. This classifier consists family of the simple probabilistic classifier.

## IV. TF-IDF AND PCFG

In this method, the performance is compared using three different sets: bi-gram frequency used by TF-IDF, PCFGs (Probabilistic context-free grammar) and combine union of both features set to infer the part of the fake news. Based on deception detection with the help of natural language processing (NLP) work done by [7] with deceptive social media reviews. The results from this review showed yield rate for predictive models which helps in classifying articles. The yield rate for the predictive model using bi-gram TF-IDF is highly effective for classification in articles from sources which are not reliable. While PCFG features have lower efficacy towards the model for classifying article from sources which are not reliable, in comparison with the findings [7] and [6], suggests that the variations obtained particularly in the classification task by using PCFG are not meaningful.

TABLE II
DIFFERENT FAKE NEWS TYPES

| | |
|---|---|
| User-based | Certain audience by fake accounts oriented towards certain audience targeting of a certain group, gender, culture |
| Visual-based | Visual based fake news which includes photo shopped images, dubbed audio or video |
| Network-based | Based on members of an organization which are interconnected, i.e. a group of friends on Facebook or individuals connected on any other social site platform |
| Post-based | Posts appeared to be on social media platforms, i.e. Facebook post with video and audio, etc. |
| Knowledge-based | Such contains a scientific explanation or reasonable issues that are not resolved spreading false information |
| Style-based | Presentation of writing style. |
| Stance-based | In Stance-based, it focuses on the construction of statements which are made in an article. The articles presented in a way to provide enough amount of details on the reader's information to provide a very little or lesser amount of information about the matter described in that subject resulting in fake arguments. |

TABLE III
COMPARISON BETWEEN TOP RELIABLE AND UNRELIABLE SOURCES

| Top 5 unreliable sources | Frequency | Top 5 reliable sources | Frequency |
|---|---|---|---|
| Before It's News | 2066 | Reuters | 3898 |
| Zero Hedge | 149 | BBC | 830 |
| Raw Story | 90 | USA Today | 824 |
| Washington Examiner | 79 | Washington Post | 820 |
| Infowars | 67 | CNN | 595 |

This explains the difference between fake news and deceptive review [6].

The methods existing on deception detection using machine learning are focused mostly on classifying reviews that are available on social media posts which are made publicly and online based reviews. The shallow POS (part-of-speech) and content related n-grams tagging have proven insufficient for classifying tasks which often fails to take consideration of important context information [8]. Only with the more complex method analysis these methods are useful [8]. Combining n-grams methods with deep syntax analysis using PCFG (Probabilistic Context-Free Grammar) have shown valuable results. The results obtained from an online review of articles have shown 85 to 91 percent accuracy in classification task using deception [7].

The implementation of a semantic analysis pairs for contradictions between text and looking at object: descriptor for additional improvement on its previous initial deep syntax model [9]. Similar results are obtained which uses vector space model for rhetorical structure [10].

### A. Performance Evaluation

The performance evaluation of trained models are evaluated on three feature sets: 1) Normalized frequency of parsed syntactical dependencies, 2) Bi-gram Term Frequency-Inverse Document Frequency and 3) Union of (1) and (2). To conduct tokenization, named entity recognition, part-of-speech and syntactical parsing the feature generation rely on Spacy Python package [11]. For fast performance, Spacy is executed in Cython, which is a another set of the Python Language that allows C code spawned from python by using the Python/C API. By implementing using Cython, performance is quick as compared to NLTK (NLP packages) [12].

In the context of speed and performance, Spacy achieves better yield rate for performance especially on tasks such as entity recognition and parsing. As compared with other tools and after evaluating peer-reviewed journals, Spacy has an advantage in the context of speed. That's why Spacy is preferred over Java applied by Stanfords PCFG (Probabilistic Context-Free Grammar) [13].

### B. Naive and Random models for baseline comparison

The baseline comparison between the two methods helps in understanding the performance. In this case, articles are from origins that are trustworthy, and the Naive Bayes model forecast all majority class. The second model is randomly selecting a classification for each article as trustworthy or not is based on the posterior probability of the class in the training set. The detail performance is shown in Table V.

### C. Combining TF-IDF bi-gram and PCFG

By combining both sets of features, the model shows the performance [6] above baseline as shown in Table VI.

The best model inclined to be stochastic gradient descent models, as they tend to perform well with high dimensional data and spare. This model out performs on precision while retaining that these products work well as recognizing the high priority articles.

### D. TF-IDF

By removing the PCFG feature allows understanding more depth of values instead of combining value.

The PCFG improves most of the metrics across the model. As PCFG predicts adds little predictive value to the models. The results [6] are shown in Table VII.

### E. PCFG only model Performance

This feature allows to isolate the predictive value, and mode result are displays. All the models give a similar result, and all model produces the same rank. Also, the threshold is switched from 0.70 thresholds to a top-k of 0.05. Because the spread of scores are low meant with a tight span. The categorization

TABLE IV
FEATURE CATEGORIZATION FOR FAKE NEWS DETECTION.

| | |
|---|---|
| Linguistic Features | Using key linguistic feature from fake news Ngrams are Stored as TFIDF values for information retrieval. Bigrams and Unigrams are obtained from a group of words in a story. TFIDF reflects the importance of a word to the document in which it is used. |
| Punctuation | It helps in differentiating between truthful and deceptive texts. This feature collects 11 types of punctuation implemented through this detection. |
| Psycho-linguistic feature | Using lexicon LIWC (Linguistic Inquiry and Word Count) for the appropriate portion of words,allowing the system to determine the language (reflecting emotions and perceptual process behind it), word counts (representing statistics of the text), a category representing part of speech (verbs, articles). This LIWC can cluster multiple sets of features: emotional tone, analytical thinking, a psychological process such as social and effective processes, linguistic processes (pronouns, and function words). |
| Syntax | Based on Context Free Grammar, set of features are extracted which depends on the lexicalized production rules which combine with their grandparent and parent nodes. For more retrieval purposes, functions from this set are encoded in TFIDF. |
| Readability | Extraction based on content features like some long words used, difficult words, characters, syllables, type of words and paragraphs. Based on this content features allows performing readability metrics like Gunning Fog, Automatic Readability Index (ARI), Flesch Reading Ease, Flesch-Kincaid. |

TABLE V
PERFORMANCE METRICS OF NAIVE AND RANDOM MODEL.

| Model | Accuracy | Precision | Recall |
|---|---|---|---|
| Naive | 67.89 percent | 54.22 percent | 54.22 percent |
| Random | 56.42 percent | 32.18 percent | 32.18 percent |

of threshold tedious and results were illuminating for 0.70. This adds task for classification, PCFG does not have a sturdy origin of information for classification by itself. The results [6] are shown in Table VIII.

## V. CNN AND RNN MODELS

### A. CNN (Convolution Neural Network)

CNN is popular for text mining and image recognition [18], [19]. By adding of the hybrid method, it would boost the performance of the model, and for content-based fake news detection, it would give better results. But for hybrid implementation text only implementation is involved so far.

### B. RNN (Recurrent Neural Network)

For time and sequence based predictions, this type of neural network is shown to get desired result [16]. As the twitter posts linked to an event which occur in time and the retweet between one user to another user is being performed within a given timeframe in a sequential mode. Depending on the variation of timeframe the rumors are examined [17]. Initially, Recurrent Neural Networks were related with the adaptation of the weights over time. The techniques are taken to solve exploding gradient was penalties, truncated backpropagation, and gradient clipping and the Echo State Networks (ESN), Long-Short Term Memory (LSTM) and dynamic weight initialization for vanishing gradient problem. For incorporating of preserving memory from the last phase for prediction of the task in the model of neural network as weights are long-term memory of the neural network.

### C. Hybrid CNN and RNN

The process [5] is divided into two parts, first by identifying automatically of features within the post on Twitter without the previous understanding about the topic of discussion using the application of a CNN and LSTM models based on hybrid deep learning model. The second approach is by classification and determination of news posts that are not real on Twitter using images and text.

Based on automatic feature extraction enabled by using deep learning models, the words that are dependent among each other can be recognized instinctively without explaining them in network expressively [15].

The deep learning architecture is implemented on three variants of deep neural network which include the models applied to train the dataset to include the items in Table IX.

The datasets [5] include 5,800 tweets which is centered on five rumor stories. The tweets were labeled as non-rumors and rumors. The events were widely reported in the social media platform. On the entire dataset of 5,800 tweets ten-fold crossvalidation was applied (i.e., for uniform inclusion in the feature vector for processing and analysis).

The hyper-parameters were improved using grid search approach, and the values are derived for the activation functions, epochs, batch size, dropout regularization (placed at 20 percent) and learning rates.

The approach of hybrid method was accompanied with the natural language semantic processing of text by allowing clarity to the posts and enhances in the identification repository. Also, in the propagation of these content on Twitter based on fake news and messages, it would be intriguing to look at the impact by the users. The aim is to find the accuracy of the posts on Twitter which can utilize in assisting law enforcement agencies to reduce the extent of the spread of such messages having negative impacts.

In finding out the posts of the news which is not real without previous understanding on the topic, the deep learning model achieves about 82 percent accuracy on the classification task . The plain vanilla LSTM model achieved 82 percent which is

TABLE VI
TF-IDF BI-GRAM AND PCFG FEATURES AT 0.7 SCORE THRESHOLD AVERAGE MODEL PERFORMANCE FOR CATEGORIZATION.

| Model | Area Under Curve (in Percent) | Precision (in Percent) | Recall(in Percent) | Accuracy (in Percent) |
|---|---|---|---|---|
| Bounded Decision Tree | 65.9 | 66.9 | 37.9 | 67.6 |
| Gradient Boosting | 75.6 | 40.2 | 16.1 | 65.7 |
| Random Forests | 80.0 | 84.2 | 18.4 | 64.8 |
| Stochastic Gradient Descent | 87.5 | 74.1 | 71.7 | 65.7 |
| Support Vector Machine | 84.3 | 80.9 | 44.5 | 73.6 |
| Baseline | - | 32.18 | 32.18 | 67.89 |

TABLE VII
TF-IDF BI-GRAM ONLY FEATURES AT 0.7 SCORE THRESHOLD AVERAGE MODEL PERFORMANCE FOR CATEGORIZATION.

| Model | Area Under Curve (in Percent) | Precision (in Percent) | Recall(in Percent) | Accuracy (in Percent) |
|---|---|---|---|---|
| Bounded Decision Tree | 60.7 | 58.5 | 23.3 | 66.1 |
| Gradient Boosting | 79.4 | 41.0 | 22.3 | 68.7 |
| Random Forests | 78.8 | 82.9 | 25.3 | 67.6 |
| Stochastic Gradient Descent | 88.3 | 88.8 | 45.3 | 77.2 |
| Support Vector Machine | 85.6 | 81.3 | 48.1 | 76.2 |
| Baseline | - | 32.18 | 32.18 | 67.89 |

TABLE VIII
PCFG ONLY FEATURES AT 0.7 SCORE THRESHOLD AVERAGE MODEL PERFORMANCE FOR CATEGORIZATION.

| Model | Area Under Curve (in Percent) | Precision (in Percent) | Recall(in Percent | Accuracy (in Percent) |
|---|---|---|---|---|
| Bounded Decision Tree | 50.0 | 40.9 | 10.8 | 60.1 |
| Gradient Boosting | 50.0 | 40.9 | 10.8 | 60.1 |
| Random Forests | 50.0 | 40.9 | 10.8 | 60.1 |
| Stochastic Gradient Descent | 50.0 | 40.9 | 10.8 | 60.1 |
| Support Vector Machine | 50.0 | 40.9 | 10.8 | 60.1 |
| Baseline | - | 32.18 | 32.18 | 67.89 |

TABLE IX
THREE VARIANTS OF DEEP NEURAL NETWORKS.

| | |
|---|---|
| LSTM | The LSTM recurrent neural network was adapted for the classification of data in sequence. For the deep learning classification method involving classification for text, the LSTM remains a popular one. |
| LSTM (with dropout regularization) | With drop regularization LSTM layers between the word embedding layer. To avoid overfitting of the training dataset LSTM was adopted. By this approach, the weights total to 20 percent of the neurons in the LSTM layer were selected randomly and dropped. |
| LSTM with CNN | After the word embedding the layer of the LSTM model an ID CNN is included. While maintaining the depth and to steer away from overfitting of the training data, further max pooling layer is added to reduce the dimensionality of the input layer.This helps in reducing resources and computational time in training the model which focuses on the overall aim to improve the accuracy of the prediction in the model. |

TABLE X
PROPOSED DEEP LEARNING METHODS IN FAKE NEWS DETECTION

| Technique | Accuracy | Precision | Recall | FMeasure |
|---|---|---|---|---|
| LSTM | 82.29 | 44.35 | 40.55 | 40.59 |
| LSTM DROP | 73.78 | 39.67 | 29.71 | 30.93 |
| LSTMCNN | 80.38 | 43.94 | 39.53 | 39.70 |

better at results regarding the recall, F-measure, and accuracy as compared to other methods.

The performance of the LSTM method with drop-out regularization was least regarding metrics which were adopted. This is due to the lack of sufficient training data, underfitting the model and examples within the network. Also, the depth of the network could be one of the reasons as it is relatively shallow that results into the drop out layer closer to the input and output layer of the model. To avoid this factor affecting the performance, an alternative is to improve it through Batch Normalization. The LSTM-CNN hybrid model performed better than the LSTM dropout regularization model with accuracy and FMeasure of 74 and 39.7 percent respectively.

## VI. COMPARISON BETWEEN ML AND DEEP LEARNING

The above methods described for detection of fake news has different aspects regarding improving performance and accuracy. Even the simple approaches such as Naive Bayes classifier showed good results on the detection of fake news. It was based on the spam filtering of email. The classification accuracy would be significantly improving by using complex model. The results obtained by using TF-IDF was potentially predictive of fake news, making an important difference in classification task for identification. In case of PCFG was good for implementation of filtering fake news as compared to the targeting fake news sites for review. TF-IDF shows potential predictive power even when named entities are ignored. But this approach would be robust to changing news cycle which requires more corpus. As TF-IDF is a vectorized approach which gives limitation and prevents generalization. The approach by using a hybrid model based on CNN and RNN is leveraged on the hybrid implementation of two deep learning models which do not require a very large number of tweets about an event to determine the credibility of the messages. This approach does not require a large amount of data typically associated with deep learning models by collecting reactions of other users based on the messages via Twitter API. This will improve better insights. The data and the method based on feature set are different but in general context for better performance on detecting fake news classification task is a must. Based on this TF-IDF and Hybrid CNN and RNN model was able to give better performance as compared to other methods and implementation represented in this survey paper.

## VII. CONCLUSION

In this paper, we studied various types of fake news and categorized the related features. Some traditional machine learning approaches, such as clustering, Naive Bayes, etc., are evaluated for detection accuracy. Some other approaches, as well as the latest deep learning approaches, such as hybrid CNN and RNN, are also evaluated to compare the performance with the traditional machine learning methods. In the future, we will further study how to select a machine learning or deep learning method for problem solving regarding the balance between accuracy and lightweightness.

## REFERENCES


[1] M. L. Della Vedova, E. Tacchini, S. Moret, G. Ballarin, M. DiPierro and L. de Alfaro, "Automatic Online Fake News Detection Combining Content and Social Signals," 22nd Conference of Open Innovations Association (FRUCT), Jyvaskyla, Finland, 2018, pp. 272 - 279.

[2] S. B. Parikh and P. K. Atrey, "Media-Rich Fake News Detection: A Survey," 2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR), Miami, FL, 2018, pp. 436 - 441.

[3] M. Granik and V. Mesyura, "Fake news detection using naive Bayes classifier," 2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON), Kiev, 2017, pp. 900 - 903.

[4] Aditi Gupta, Hemank Lamba, Ponnurangam Kumaraguru, and Anupam Joshi, "Faking Sandy: characterizing and identifying fake images on Twitter during Hurricane Sandy," In Proceedings of the 22nd International Conference on World Wide Web (WWW'13 Companion). ACM, New York, NY, USA, 2013, pp. 729 - 736.

[5] Oluwaseun Ajao, Deepayan Bhowmik, and Shahrzad Zargari, "Fake News Identification on Twitter with Hybrid CNN and RNN Models," In Proceedings of the 9th International Conference on Social Media and Society (SMSociety'18). ACM, New York, NY, USA, 2018, pp. 226 - 230.

[6] S. Gilda, "Evaluating machine learning algorithms for fake news detection," 2017 IEEE 15th Student Conference on Research and Development (SCOReD), Putrajaya, 2017, pp. 110 - 115.

[7] S. Feng, R. Banerjee, and Y. Choi, Syntactic stylometry for deception detection, In Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Short Papers-Volume 2, Association for Computational Linguistics, 2012, pp. 171 - 175.

[8] N. J. Conroy, V. L. Rubin, and Y. Chen, Automatic deception detection: Methods for finding fake news, Proceedings of the Association for Information Science and Technology, vol. 52, no. 1, pp. 1 - 4, 2015.

[9] V. W. Feng and G. Hirst, Detecting deceptive opinions with profile compatibility, in IJCNLP, 2013, pp. 338 - 346.

[10] V. L. Rubin and T. Lukoianova, Truth and deception at the rhetorical structure level, Journal of the Association for Information Science and Technology, vol. 66, no. 5, pp. 905 - 917, 2015.

[11] Explosion, Spacy, Sep. 2017. [Online]. Available: https://github.com/explosion/spaCy.

[12] S. Bird, E. Klein, and E. Loper, "Natural language processing with Python: analyzing text with the natural language toolkit," OReilly Media, Inc., 2009.

[13] J. D. Choi, J. R. Tetreault, and A. Stent, It depends: Dependency parser comparison using a web-based evaluation tool., in ACL (1), 2015, pp. 387 - 396.

[14] L. Buitinck, G. Louppe, M. Blondel, F. Pedregosa, A. Mueller, O. Grisel, V. Niculae, P. Prettenhofer, A. Gramfort, J. Grobler, R. Layton, J. VanderPlas, A. Joly, B. Holt, and G. Varoquaux, API design for machine learning software: Experiences from the scikitlearn project, in ECML PKDD Workshop: Languages for Data Mining and Machine Learning, 2013, pp. 108 - 122.

[15] Jing Ma, Wei Gao, Prasenjit Mitra, Sejeong Kwon, Bernard J Jansen, KamFai Wong, and Meeyoung Cha, "Detecting Rumors from Microblogs with Recurrent Neural Networks," In IJCAI. 2016, pp. 3818 - 3824.

[16] Jing Ma, Wei Gao, Zhongyu Wei, Yueming Lu, and Kam-Fai Wong, "Detect rumors using time series of social context information on microblogging websites," In Proceedings of the 24th ACM International Conference on Information and Knowledge Management. ACM, 2015, pp. 1751 - 1754.

[17] Sejeong Kwon, Meeyoung Cha, and Kyomin Jung, "Rumor detection over varying time windows, " PLoS One. 2017; 12(1): e0168344.

[18] Shiou Tian Hsu, Changsung Moon, Paul Jones, and Nagiza Samatova, "A Hybrid CNN-RNN Alignment Model for Phrase-Aware Sentence Classification," In Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers, 2017, Vol. 2. pp. 443 - 449.

[19] Connor O'Donnell and Wenlin Han, "Recent Developments and Emerging Challenges in Deep Neural Networks for Cybersecurity and Other Applications," In the Proceedings of the 20th International Conference on Artificial Intelligence (ICAI'18), 2018, pp. 112 - 118.