IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-7, Issue-4, April 2019 E-ISSN: 2347-2693

A Hybrid Approach for Fake News Detection using Convolution and Multilayer Perceptron

Mohd Zeeshan Ansari[1], Mumtaz Ahmed[2*]

[1,2]Dept. of Computer Engineering, Jamia Millia Islamia, New Delhi, India

Corresponding Author: mahmed1@jmi.ac.in

DOI: https://doi.org/10.26438/ijcse/v7i4.11811187 | Available online at: www.ijcseonline.org

Accepted: 19/Apr/2019, Published: 30/Apr/2019

***Abstract*—** Social media platforms allow its users to publicly share any kind of content without any restriction. This shared content is available to a very large number of people having access to social media, moreover, it plays a significant role in casting their trust and belief. Due to this, there is an essential necessity to probe the genuineness and authenticity of the publicly shared content. Fake news is one such problem which has recently attracted enormous attention due to its large social, political and economic impacts on an individual and the society. Manual analysis of articles on social media is a cumbersome task and also it does not ensure a high success rate in the detection of fake news. In this article, we proposed a hybrid deep learning architecture to exploit the characteristics of Convolutional Neural Network along with Multilayer Perceptron. To evaluate the architecture, we used LIAR dataset which contains the news text and profile of the news source. After testing the architecture on various models a significant improvement was observed when compared to state of the art models.

***Keywords*—** Fake News, Deception, Convolutional neural network, Multilayer perceptron.

I. INTRODUCTION

The rapid growth in the users of social media platforms have introduced several typical problems in its text analysis in all domains of natural language processing. A user of a social networking site or a blog is open to share his or her choice of content which is predominantly observable to a very large set of users. The principle behind spread of information is established through the need that satisfies towards the people's interest and fits their system of belief [1]. Nowadays, social media is very popular and desirable source of news as compared to electronic media due to ease of accessibility and dissemination. A principally noticeable observation is that the readers most often don't pay attention towards the reliability of original source of generation as well as the authenticity of publicly shared content [2]. Therefore, an essential necessity arises to automatically probe the genuineness of publicly shared content, moreover it may be considered as an extremely important and an open challenging problem. *Fake news* is such kind of problem in which misinformation is intentionally spread for various purposes. It has gained popularity recently, due to its social, political and economic impact on the individual and society.

The heterogeneous and complicated nature of fake news creates large dissimilarity over its definition. It can be broadly classified into three categories: (*i*) intentionally generated news, (*ii*) satirical news (*iii*) deceptive or distorted news [3]. Intentionally generated news is simple and broadly accepted definition which is defined as a misrepresentation of news in order to mislead targeted readers [4-6]. Satirical news is generated for entertainment nonetheless, its content are false leading to certain kind of frauds [7]. Deceptive or distorted news most often includes severe deceptions, satires and hoaxes. [9-12].Moreover, Conroy (2017) gave the definition of *fake news detection* as "the task of categorizing news along a continuum of veracity with an associated measure of certainty" [13].In a study, Rubin(2017) concluded that on the basis of human assessment, the success rate of identifying a fake news is only 50-63 % [14]. As the analysis of its contents is fairly wide and complex in nature, it has attracted considerable attention by natural language scientists and computational linguistics and psychologists [15-17]. Subsequently, the electronic media platforms have initiated collaboration with professional journalists and machine learning scientists in order to build fake news detection systems. Such systems must be able to identify and generate automatic alerts on spread of fake contents over the internet. Wu and Liu (2018) studied information diffusion on fake news data and proposed a framework to trace the similar sources of fake news generation [18].

The analysis of fake news extremely challenging task and its detection on social media and other online platforms has become the active domain in research. It seems impossible to stop the spread of Fake News as with speed and scale it is spreading. Several machine learning and deep learning approaches are explored for fake news detection [19-23]. However, the state of the art models also do not yield significant results on the newer challenges [3]. In this article, we have presented a hybrid architecture for fake news detection which is composed if Convolutional Neural Networks and Multilayer Perceptron. The Fake News Challenge Dataset is employed the proposed architecture. It is observed that the proposed architecture outperforms the state of the art developments in fake news detection. The organization of our work in the article is as follows. The section II presents the related work, the section III presents the overall architecture of the model. The section IV presents the dataset statistics and results along with its analysis. The last section gives the conclusion of the work.

II. Related Work

In literature, various methods have been proposed for automatic detection of Fake News on social media platforms. Several works are found on Fake News detection using machine learning and modern deep learning methods. The simple Naïve Bayes classifier with Bag of Words language model was initially employed to identify the Fake News and its dataset was collected for this experiment from the Facebook's posts [24]. Another work examines the Term Frequency-Inverted Document Frequency and n-gram techniques as the feature extraction methods and applies six machine learning algorithms to classify the news. These features along with linear support vector machine performed we as compared to other algorithms [25]. In 2007, the fake news challenge (FNC-1) task held over internet drew considerable attention of researchers towards this problem. Several neural network and deep learning models were submitted in the task. Subsequently, a new dataset was published for fake news detection which contain 12886 labelled statements in different contexts. The dataset collected from Politifact.com, also put forward the trusted analysis report and links to source documents. The wide range of methods for fake news detection and deception detection are tested on this dataset.

An attention based LSTM model is proposed to integrate speaker profiles into LSTM model for fake news detection. This hybrid model is made up by concatenate two LSTM models in which the speaker profile such credit history, speaker, speaker's job, state and party affiliation are used in attention model. The speaker profile is used for topic information of the news clauses then both the LSTM models are concatenated, a soft-max activation function is used before final classification. The results show that the hybrid model has improved the accuracy by 14.5% on LIAR: a benchmark dataset for Fake News detection [26]. A different approach for stance detection is used by employing various LSTM attention models such as conditionally encoded LSTM, global attention based LSTM, word by word attention based LSTM, bidirectional global attention and bidirectional conditional LSTM with bidirectional global attention [10]. Wang (2017) presented hybrid CNN model for fake news and reported significant improvements in performance as compared to logistic regression and support vector machines. Two different variations of CNN are constructed in order to build a mixture model in which one model is CNN on the news text data whereas the second model is CNN on news meta-data. [27].A recent work on deception detection concentrates on the deception detection on user's review taken from online shopping[28-29] sites. The stacking ensemble methods for fake news detection using FNC-1 are being employed and significant improvements have been reported [30].

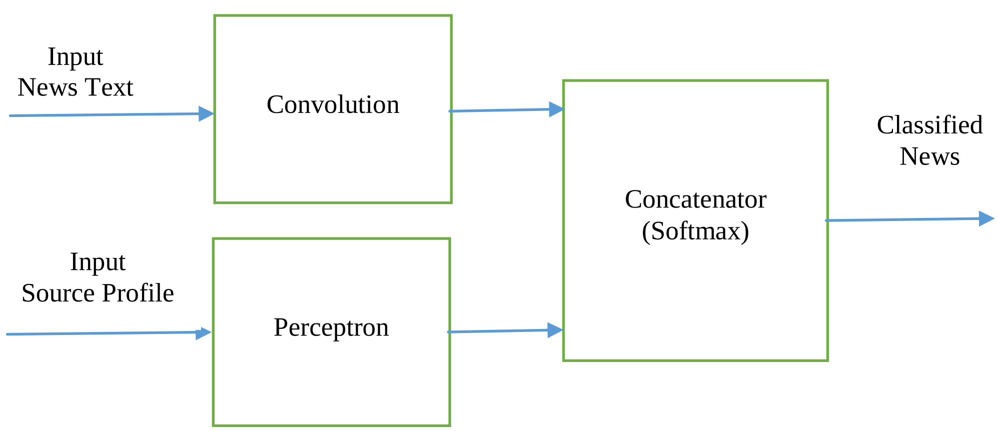

Fig. 1. Hybrid Architecture for Fake News Detection

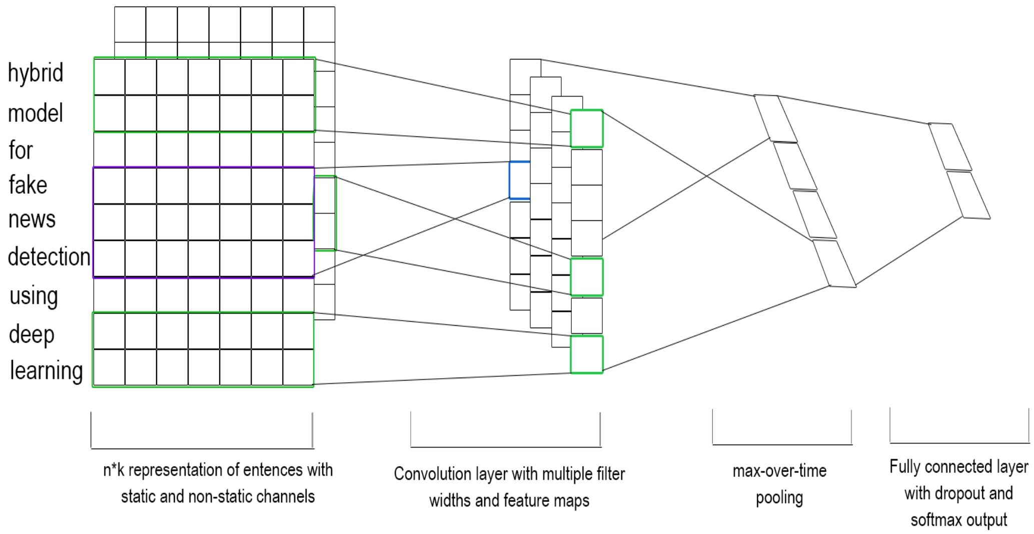

Fig. 2. Text based Convolution for Fake News articles

III. THE HYBRID ARCHITECTURE USING CNN AND MLP FOR FAKE NEWS

The traditional fake news detection totally relies upon the news content whereas in social media the user profile can be considered as a useful auxiliary information. Therefore we present an architecture that employs separate models for both the news content and the speaker profile. The characteristic of convolutional networks to learn the features efficiently fascinated us to apply it in this problem. The primary objective of this architecture is to improve fake news classification performance by exploring the mixture of convolutional neural models and perceptron models on news content and profile respectively. The Fig. 1shows the three major components of the architecture: the news text model, the profile model and the concatenator. The news text article is transformed into word vectors generated using Glove [31]. The convolution applied to these vectors based on the model by Kim (2010) which is a deep model that enables the convolution operations on the text of the news article and subsequently, generates the feature maps of news articles[32]. The multilayer perceptron takes the speaker profile as input and generates the features for metadata. The concatenator is the soft max function which generates the final classified output.

A. Fake news

A news object can be significantly separated into the article text and the speaker profile. For a given $\boldsymbol{N}$ news objects, let n_k be each news, for $k = 1$ to $\boldsymbol{N}$. Let t_k and S_k be the collection of $\boldsymbol{N}$ news articles and collection of speaker $\boldsymbol{N}$ profiles respectively. Each news article text t_k consists a sequence of words, $w_{k1}, w_{k2}, w_{k3....}w_{kl},$ where l is the length of that article. It which may be specified as $t_{k=}\ w_{11} \oplus w_{12} \oplus ..... \oplus w_{1l},$ where, $\oplus$ is a concatenation operator. Each word may be represented by the vector of length L. The speaker profile S_k is composed of speaker name S_S, party affiliation S_P, job S_J, credit history S_H and subject topic S_T. In order to build the hybrid architecture, news article t_k is input to the text convolutional model and the corresponding speakers profile S_k are input to the multilayer perceptron model. Finally, the outputs of both the models are then concatenated with the help of soft-max activation function for classification. The whole model is characterized into three separate models: (1) text convolutional model of news articles, (2) multilayer perceptron model of speaker profile, and (3) softmax concatenation of both the models.

B. Features

The description of features present in the speaker profile which are significant for consideration in the proposed model are:

(*i*) **History.** Old records of truth and untruth statements or news articles posted on internet.
(*ii*) **Speaker.** Name of the speaker.
(*iii)* **Subject.**Topic of news viz. political news, sports news,etc.

Table 1. LIAR Dataset statistics.

	Number of words
Train set	10269
Validation set	1284
Test set	1283
Average size of article	18
Total size of dataset	12836

Table 2. Class-wise Distribution of LIAR tagset

tagset	%	Number of entities
Pants fire	8.18	1050
False	19.56	2511
Barely-true	16.42	2108
Half-true	20.55	2638
Mostly-true	19.21	2466
True	16.08	2063
Total	100	12836

(*iv*) **Party.** Organization to which speaker belongs.

(*v*) **Job.** Work profile or designation of the speaker.

C. *The News Text Model*

Embedding layer. In text based convolutional model, the embedding layer encodes each word of every news article in the form of vector of specified size. We trained the vectors on different sizes of 50, 100, 200 and 300 separately to evaluate the effect of vector size on the model. Since the training of the network requires the samples in batches, each of the sample vector is appended with zeros in order to obtain same length of all the samples. In this way, the first layer i.e. the embedding layer of CNN is constructed as a two dimensional matrix of text. We employed pre-trained GloVe word vectors to construct this embedding layer [31].

Convolution layer. The output of embedding layer is matrix of size 57 X100is the vector dimension. The convolutional filter is applied to each possible window of words of the matrix in different filter size n=2,3,4 and 5 which results into feature maps. A feature $C_i$is obtained from a window of wordst_{k+x-1} by

$$C_i = f(F.t_{k+x-1} + b) \quad (1)$$

where,$b$is bias factor, f is the non-linear activation function and F is the convolution filter which is applied to window of words in the text to produce feature maps.By this process each filter generates a single feature map. We can use multiple filters to produced large number of features from same window size filters. We use bi-gram, tri-gram, four-gram, five-grams as filters which identify the context of words being used.

Pooling layer. The max-pool operation is applied on the feature map to select the maximum value as the feature corresponding to this filter. Max-pool operation reduces the feature map by taking in consideration only maximum values from the feature map. After all the features are extracted, a flatten layer is used followed by a dense layer.

D. *The News Metadata Model using MLP*

The input layer of multilayer perceptron model has the number of nodes equal to the number of features selected from speaker profile. The two hidden layers are used in MLP having 100 and 180 hidden nodes followed by a dense layer. Both the models use back-propagation algorithm for training.

E. *Concatenator*

Both convolutional and perceptron models are concatenated using the soft-max activation function as a non-linear activation function before final classification of news articles.

Table 3. Comparative analysis of proposed, baseline and existing models.

Models	Validation Accuracy	Test Accuracy
Proposed CNN-MLP Hybrid Model	**0.436**	**0.424**
$\text{LSTM}_{[26]}$	0.407	0.415
$\text{CNN}_{[27]}$	0.277	0.274
$\text{CNN}_{[32]}$	0.260	0.270
Random Forest	0.265	0.263
SVM	0.258	0.255
Logistic Regression	0.257	0.247
Bi-LSTM	0.223	0.233

Table 4. Performance of proposed model on several combination of Speaker profile attributes.

Text + Attributes of Speaker Profile	Validation Accuracy (%)	Test Accuracy (%)
Text + Speaker + History	**43.66**	**42.43**
Text +Subject +History	43.40	41.19
Text +Subject + Speaker +Party +History	43.61	41.66
Text + Subject +Party+ History	43.38	41.80
Text + Party + Job + History	44.16	40.90
Text +Party + History	44.37	40.51
Text + Job + History	42.76	41.84
Text +Subject + Speaker +History	43.69	40.25
Text +Subject +Job+ History	42.45	41.20
Text +Subject + Speaker +Party	42.35	41.20
Text +Subject + Speaker+ History	42.29	40.75
Text +Speaker +Job+ History	42.99	39.94
Text +Speaker +Party+ History	43.30	39.20
Text +Subject + Party + History	41.82	40.24
Text + Speaker + Party + Job +History	41.51	40.48
Text +Subject + Speaker +Job + History	42.06	39.47

IV. RESULTS AND DISCUSSION

The proposed hybrid architecture has been evaluated using the benchmark LIAR dataset [27]. Before feeding the raw text into network, considerable amount of pre-processing is performed. Firstly, the unidentified characters were removed from the text.Secondly, the sentences were tokenized into words.Finally the feature vector is normalized,moreover, as a necessity for each news article to have same length, hence each article is padded with null words, if required.The pre-processed dataset is divided into training, validation and testing samples as shown in Table 1.

A. LIAR Dataset

LIAR is a high magnitude dataset as compared to prior datasetsdue to which it enables the advancements of statistical and computational methods for fake news detection. The dataset includes total 12836 instances on 141 topics from *politifact.com*, which maintains the comprehensive report of news statements and their links to source articles. Each instance includes the text that is the content news article and corresponding meta-data, which is, new id, speaker profile and class. Total seven attributes in Speaker profile are as speaker name, party affiliation, speaker's job title, subject, context, the state info and credit history of speaker. The credit history of speakers profile includes the past record of inaccurate statements for each speaker. There are six labels in the dataset are as pants on fire, false, barely-true, half-true, mostly-true and true and their distribution is shown in Table 2.We evaluated the proposed hybrid CNN and MLP model for fake news detection on LIAR dataset.

B. Experimental on LIAR Dataset

In the experiment, we used mixture of CNN and MLP to construct the proposed hybrid model. CNNis applied on the text of news articles consists of the embedding layer, the convolution layer followed by the pooling layer.The embedding layer is constructed using the glove word-vectors ofsize10dimensions to represent words into vector space. In this experiment, we employed three filters of bi-gram, tri-gram and four-gram which extract 512 features for each size of filter. A zero dropout layer is used to reduce irrelevant training parameters in the model. On the other hand, the multilayer perceptron applied on speaker profile has input layer, two hidden layers and an output layer.The relu activation function is used as non-linear activation function is used after each layer in both CNN and MLP. The overall modelis trained for 25 epochs with batch size 16. Model converged after 20 epochs and training and testing loss became constant.

C. Analysis of Results

Our proposed model shows better accuracy against the equivalent deep learning models. It is also observed that our model shows considerably better performance when compared to various machine learning models. The comparative performance of models is shown in Table 3 which includes attention based LSTM [26], hybrid CNN [27], CNN [26], random forest, support vector machine and regularized logistic regression classifier. The comparative performance of models is shown in Table 3, which includes attention based LSTM [26], hybrid CNN [27], CNN [26],

Table 5. Performance of proposed model on several combination of Speaker profile attributes.

Text + Attributes of Speaker Profile	Validation Accuracy (%)	Test Accuracy (%)
Text + Speaker + History	**43.66**	**42.43**
Text +Subject +History	43.40	41.19
Text +Subject + Speaker +Party	43.61	41.66
Text + Subject +Party+ History	43.38	41.80
Text + Party + Job + History	44.16	40.90
Text +Party + History	44.37	40.51
Text + Job + History	42.76	41.84
Text +Subject + Speaker +History	43.69	40.25
Text +Subject +Job+ History	42.45	41.20
Text +Subject + Speaker +Party	42.35	41.20
Text +Subject + Speaker+ History	42.29	40.75
Text +Speaker +Job+ History	42.99	39.94
Text +Speaker +Party+ History	43.30	39.20
Text +Subject + Party + History	41.82	40.24
Text + Speaker + Party + Job +History	41.51	40.48
Text +Subject + Speaker +Job +	42.06	39.47

random forest, support vector machine and regularized logistic regression classifier.

Wang (2017) adopted hybrid CNN for fake news detection and reported that I t performs better than text only CNN model. We evaluated this model in LIAR dataset and observed the accuracy of around 27.7 %. We also evaluated the LIAR dataset on random forest algorithm with n-gram and TF-IDF feature extraction methods and observed the testing accuracy of nearly 26.3%.

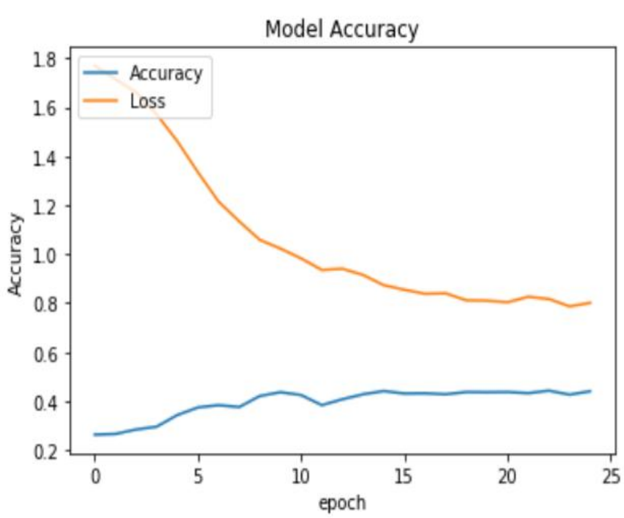

Fig. 3. Validation accuracy and loss with respect to epochs in case of Text along with Credit History

The proposed model is evaluated on a variety of configurations for the attributes of speaker profile displayed in Table.3. Best performance is observed in Text+Speaker+History as compared with different attribute combinations. The validation accuracy and test accuracy obtained is 43.66% and 42.43% respectively. The validation and test accuracy of our model for all the metadata combinations is shown in Table 3.

V. CONCLUSION

The proposed work is a mixture model of Convolutional Neural Network and Multilayer Perceptron for detection of fake news over the social media platforms which considered various characteristics of source profile and textual comments about the source of news. Specifically, we considered the content of news as well as the metadata of news from the LIAR dataset. The results show the speaker profile especially the history significantly contributes to the improvement in accuracy. Different combinations of filters with word vectors are used to generate the features as input to the convolutional neural network. The experimental results on the LIAR dataset show that the proposed mixture model improved the validation accuracy by 7.13% and testing accuracy by 2.17% as compared to the latest state of the art model. The proposed model is fully dependent on the metadata of news therefore in future, this model can be transformed into new one which will have less dependent on metadata, by comparing the news article to at real time with various news sources. We can improve this model by taking into consideration different variations of Convolutional Neural Networks.

REFERENCES

[1] M.D. Vicario, A.Bessi, F.Zollo, F.Petroni, A.Scala, G.Caldarelli, H. E. Stanley, and W.Quattrociocchi. "*The spreading of misinformation online*". Proceedings of the National Academy of Sciences 113, 3, pp. **554–559**, **2016**.

[2] R.Marchi, "*With facebook, blogs, and fake news, teens reject journalistic objectivity*", Journal of Communication Inquiry, Vol.**36(3)**, pp. **246-262**, **2012**.

[3] K.Shu, A.Sliva, S.Wang, J.Tang, and H.Liu, "*Fake News Detection on Social Media: A Data Mining Perspective",* ACM SIGKDD Explorations Newsletter Vol.**19-1**, pp.**22–36**, **2017**.

[4] E. Mustafaraj and P.T.Metaxas. "*Thefake news spreading plague: Was it preventable?*" arXiv preprint arXiv:1703.06988, 2017.

[5] M.Potthast, J.Kiesel, K.Reinartz, J.Bevendor, and B.Stein, "*A stylometric inquiryinto hyperpartisan and fake news*". arXiv preprintarXiv:1702.05638, 2017.

[6] David O Klein and Joshua R Wueller. Fake news: A legal perspective. 2017.

[7] M. Balmas, "*When fake news becomes real: Combinedexposure to multiple news sources and politicalattitudes of inefficacy, alienation, and cynicism*". CommunicationResearch, Vol.**41(3)**, pp. **430-454**, **2014**.

[8] V.L Rubin, N.J. Conroy, Y.Chen, and S.Cornwell, "*Fake news or truth? Using satiricalcues to detect potentially misleading news*". In Proceedingsof NAACL-HLT, pp. **7-17**, **2016**.

[9] S.Chopra and S.Jain, "*Towards automatic identification of fake news: Headline-article stance detection with LSTM attention models*," 2017.

[10] S.Chopra, S.Jain, J.M. Sholar, "*Towards Identification of Fake News: Headline, Article Stance Detection With LSTM Attention Models*", Stanford CS224 Deep Learning For NLP Final Project, 2017.

[11] E.Fitzpatrick, J.Bachenko, and T.Fornaciari, "*Automatic Detection of Verbal Deception. SynthesisLectures on Human Language Technologies*" Morgan & Claypool Publishers.

[12] P.Rosso and L.Cagnina, "*Deception Detectionand Opinion Spam. In: A Practical Guide to SentimentAnalysis*", Cambria, E., Das, D., Bandyopadhyay,S., Feraco, A. (Eds.), Socio-Affective Computing,Vol. **5**, Springer-Verlag pp. 155-171, **2017**.

[13] N.J.Conroy, V.L.Rubin, Y.Chen "*Automatic Deception Detection*" Proceedings of the 78th ASIS&T, **2015**.

[14] V.L.Rubin and T.Lukoianova "*Truth anddeception at the rhetorical structure level*". Journal of the Association for Information Science and Technology, Vol. **66(5)**, pp. **905-917**, **2015**.

[15] A.A.Memon, A.Vrij, and R.Bull, "*Psychology and law: Truthfulness,accuracy and credibility*". John Wiley & Sons, 2003.

[16] S. R. Maier, "*Accuracy Matters: A Cross-Market Assessment of Newspaper Error and Credibility*," Journalism & Mass Communication Quarterly, Vol. **82**, Issue. **3**, pp. **533–551**, **2005**.

[17] B.J.Fogg and H.Tseng. "*The Elements of Computer Credibility*" , in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '99. New York, NY, USA: ACM, pp. **80–87**,**1999**.

[18] L.Wu, H.Liu, "*Tracing fake-news footprints: Characterizing social media messages by how they propagate*", Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining. pp. **637–645**. WSDM '18, ACM, **2018**.

[19] N.J.Conroy, V.L.Rubin and Y.Chen, "*Automatic deceptiondetection: Methods for finding fake news*". Proceedings of the Association for Information Science and Technology 52, 1 , pp. **1–4**,**2015**.

[20] Z.Jin, J.Cao, Y.Zhang, J.Zhou and Q.Tian, "*Novel visual and statistical image features for micro blogs news verification*", IEEE transactions on multimedia Vol.**19**, Issue.**3** , pp. **598–608**, **2017**.

[21] E.Tacchini, G.Ballarin, M.L.Vedova, S.Moret, and L.de Alfaro, " *Some like it hoax: Automated fake news detection in social networks*", arXiv preprint arXiv:1704.07506 ,2017.

[22] J.Ma, W.Gao, P.Mitra, S.Kwon, B. J.Jansen, K.Wong, and M.Cha. "*Detecting Rumors from Microblogs withRecurrent Neural Networks*", In IJCAI, pp **3818–3824**, **2016**.

[23] N.Ruchansky, S.Seo, and Y.Liu, "*CSI:AHybrid Deep Model for Fake News Detection*", Proceedings of the 2017 ACMon Conference on Informationand Knowledge Management. ACM, pp. **797–806**, **2017**.

[24] M.Granik, V.Mesyura, "*Fake News Detection Using Naive Bayes Classifier*", First Ukraine Conference on Electrical and Computer Engineering, IEEE conference, Ukraine, pp. **900-903**, **2017**.

[25] H.Ahmed, I.Traore, S.Saad, "*Detection of Online Fake News Using N-Gram Analysis and Machine Learning Techniques*", Traore I., Woungang I., Awad A. (eds) Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments. ISDDC, Lecture Notes in Computer Science, Vol.**10618**. Springer, Cham, **2015**.

[26] Y.Long, Q.Lu, R.Xiang, M.Li, C.R.Huang. "*Fake News Detection Through Multi-Perspective Speaker Profiles*", International Joint Conference on Natural Language Processing, AFNLP, Taiwan, pp. **252–256**, **2017**.

[27] W.Y.Wang. "*liar pants on fire: A new benchmark dataset for fake news detection*", arXiv preprint arXiv1705,00648, 2017.

[28] R.Mihalcea, C.Strapparava. "*The lie detector: Explorations in the automatic recognition of deceptive language*", Proceedings of the ACL-IJCNLP, **2009**.

[29] M.Ott, Y.Choi, C.Cardie, and J.T.Hancock' "*Finding deceptive spam by any stretch of the imagination*", Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, Vol**1**. Association for Computational Linguistics, pages **309–319**, **2011**.

[30] J.Thorne et.al. "*Fake News Detection using Stacked Ensemble of Classifiers*" , Proceedings of the EMNLP, pages 80–83Copenhagen, Denmark, September-7, Association for Computational Linguistics, **2017**.

[31] J.Pennington, R.Socher, C.D.Manning. "*GloVe: Global Vectors for Word Representation*", Conference on Empirical Methods in Natural Language Processing (EMNLP), pp. **1532–1543**, **2014**.

[32] Y.Kim. "*Convolutional neural networks for sentence classification*", Conference on Empirical Methods in Natural Language Processing (EMNLP), **2014**.

Authors Profile

Mr. Mohd. Zeeshan Ansari is currently Assistant Professor at Department of Computer Engineering, Jamia Millia Islamia (A Central University), New Delhi. He received M.Tech in Computer Science and Engineering from Delhi Technological University, New Delhi in year 2014 and B.Tech in Computer Science and Engineering from Uttar Pradesh Technical University, Lucknow, Uttar Pradesh in year 2005. He has more than ten years of teaching experience. His area of research interest is Code Mixing, Information Extraction and Retrieval, Text Mining, Natural Language Processing and Soft Computing Techniques. His field of Specialization is Information Extraction and Retrieval.

Mr. Mumtaz Ahmed pursed Bachelor of Technology in Computer Engineering from Jamia Millia Islamia, New Delhi in 2004 and Master of Engineering in Computer Engineering from Delhi Technological University New Delhi in 2013. He is currently pursuing Ph.D. and currently working as Assistant Professor in Department of Computer Engineering, Jamia Millia Islamia, New Delhi since 2007. His main research work focuses on Computer Networks, Network Security and Artificial Intelligence. He has around 10 years of teaching experience and 5 years of research experience.

