**An Intelligent Feature Selection-based Fake News Detection Model for Pandemic Situation with** **Optimal Multiscale Convolutional Neural-Long Short-Term Memory**

**Introduction**

Fake news, is refers to stories that are intentionally and verbally false is deliberately created to mislead people for financial or political gains and has existed for a long time, even before the appearance of traditional media such as the printing press [9]. Social media platforms such as Twitter or Face book and their increasing popularity speed up the dissemination of fake news since news can quickly and freely circulate through a huge network of social media users, where everyone can view and share news without paying much attention to the veracity of each reported claim [10]. The global outbreak of Corona Virus Disease 2019 (COVID-19) in the age of social media has caused millions of people to contract the virus, and a significant number of them have lost their lives, resulting in a tremendous social and economic shock across the globe [11]. Amid the growing burden of the COVID-19 pandemic, the parallel emergencies infodemic (information + epidemic) must be simultaneously tackled: the proliferation of fake news, false rumors, and misinformation surrounding COVID-19. At its worst, some health-related fake news has introduced people to ineffective or even potentially harmful remedies, which seriously disrupt the social order [12]. Early works in fake news detection are mainly based on fact-checking of external sources or the writing style of news content. The fake news detection task is traditionally approached using linguistic features that are able to identify linguistic patterns of the text [13]. The main limitation of such methods is that they are hand-crafted and involve manual labor for designing them. On the other hand, more recent deep neural networks have been proposed to alleviate the need for manually designing hand-crafted features since deep learning methods are able to automatically capture linguistic patterns [14].

As manual fact checking is a very tedious task, automatically identification of fake news has drawn considerable attention in the Natural Language Processing (NLP) community to help alleviate the burdensome and time-consuming human activity of fact checking [15]. Despite that, the task of evaluating the authenticity of news remains very complex even for automated systems [16]. Identifying fake news articles by understanding what other news organizations are reporting about the same topic could be a valuable first step. This step is known as Stance detection [17]. Stance detection has always been an important foundation for various tasks, such as analyzing online debates, determining the authenticity of rumors on twitter, or understanding the argumentative structure of persuasive essays [18]. Automatically detecting fake news has been garnering an increasing number of active research interests from the academic and industrial communities [19]. Traditional approaches extract handcrafted features from news textual content relying on expert knowledge, followed by traditional machine learning algorithms for training fake news classifier. These handcrafted features based methods are simple but lack comprehensiveness and flexibility [20]. Substantial researches have indicated that how to design artificial features is importantly critical to kinds of natural language processing tasks [21].

With the rapid advances in Artificial Intelligence (AI), a significant number of experiments are being undertaken to tackle issues that were never addressed in the framework of computer science, such as fake news detection [22]. Automatic detection approaches based on machine learning have been studied to combat the emergence and dissemination of false news [23]. The majority of fake news detection systems utilize machine learning approaches to help consumers in filtering the content they are seeing and determining if a given news piece is misleading or not. Deep learning techniques recent accomplishments in difcult natural language processing tasks make them viable for detecting fake news effectively and efficiently [24]. Creating automatic, trustworthy, and accurate systems for identifying fake news on social media is a trending topic of research. The process of determining if a certain news item on any field, from any social media domain, is purposefully or inadvertently misleading might be characterized as fake news detection [25]. Convolutional Neural Network (CNN) has been prominent in many fields with the best performance, including computer vision, smart building structures, and natural language processing. CNN uses convolution layers, pooling layers, and fully connected layers to extract more features with high-level and low-level features.

**Related works**

In 2021, Yuan *et al.* [1] have developed an approach which we term “domain-adversarial and graph-attention neural network” (DAGA-NN) model to address the challenge. Its main advantage was that, in a text environment with multiple events/domains, only partial domain sample data are needed to train the model to achieve accurate cross-domain fake news detection in those domains with few samples, which makes up for the limitations of traditional machine learning in fake news detection tasks due to news content evolution or cross-domain identification. Extensive experiments were conducted on two multimedia datasets of Twitter and Weibo, and the results showed that the proposed model was very effective in detecting fake news across events/domains.

In 2021, Silva *et al.* [2] have developed this work proposes Propagation2Vec, a novel fake news early detection technique, which assigns varying levels of importance for the nodes and cascaded in propagation networks, and reconstructed the knowledge of complete propagation networks based on their partial propagation networks at an early detection deadline. Proposed experiments have shown that developed model can achieve state-of-the-art performance while only having access to the early stage propagation networks. Furthermore, devise general explanations for the underlying logic of Propagation2Vec based on its attention weights assigned to different nodes and cascades, which improves the applicability of our approach and facilitates future research on propagation network-based fake news detection.

In 2021, Zeng *et al.* [3] have suggested a novel approach for Fake News Detection by comprehensively mining the Semantic Correlations between Text content and Images attached (FND-SCTI). First, learn image representations via the pretrained VGG model, and use them to enhance the learning of text representation via hierarchical attention mechanism. Second, a multimodal variational autoencoder was exploited to learn a fused representation of textual and visual content. Third, the image-enhanced text representation and the multimodal fusion eigenvector are combined to train the fake news detector. Experimental results on two real-world fake news datasets, Twitter and Weibo, demonstrate that developed model outperforms seven competitive approaches, and were able to capture the semantic correlations among multimodal contents.

In 2021, Saleh *et al.* [4] have recommended a novel approaches based on Machine Learning and Deep Learning for the fake news detection system to address this phenomenon. The main aim of this paper was to find the optimal model that obtained high accuracy performance. Therefore, proposed an Optimized Convolutional Neural Network Model to Detect Fake News (OPCNN-FAKE). On compareing the performance of the OPCNN-FAKE with Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), and The six regular ML techniques: Decision Tree (DT), logistic Regression (LR), K Nearest Neighbor (KNN), Random Forest (RF), Support Vector Machine (SVM), and Naive Bayes (NB) using four fake news benchmark datasets. Grid search and hyperopt optimization techniques have been used to optimize the parameters of ML and DL, respectively. In addition, N-gram and Term Frequency—Inverse Document Frequency (TF-IDF) have been used to extract features from the benchmark datasets for regular ML, while Glove word embedding has been used to represent features as a feature matrix for DL models. To evaluate the performance of the OPCNN-FAKE, accuracy, precision, recall, F1-measure were applied to validate the results. The results have shown that OPCNN-FAKE model has achieved the best performance for each dataset compared with other models. Furthermore, the OPCNN-FAKE has a higher performance of cross-validation results and testing results over the other models, which indicated that the OPCNN-FAKE for fake news detection was significantly better than the other models.

In 2020, Umer *et al.* [5] have initiated a hybrid Neural Network architecture, that combines the capabilities of CNN and LSTM, was used with two different dimensionality reduction approaches, Principle Component Analysis (PCA) and Chi-Square. This work proposed to employ the dimensionality reduction techniques to reduce the dimensionality of the feature vectors before passing them to the classifier. To develop the reasoning, this work acquired a dataset from the Fake News Challenges (FNC) website which has four types of stances: agree, disagree, discuss, and unrelated. The nonlinear features are fed to PCA and chi-square which provides more contextual features for fake news detection. The motivation of this research was to determine the relative stance of a news article towards its headline. The proposed model improved results in terms of Accuracy and F1−score. The experimental results show that PCA outperforms than Chi-square and state-of-the-art methods with 97.8% accuracy.

In 2022, Wang *et al.* [6] have aims to reveal the factors influencing the acceptance of fake news rebuttals on Sina Weibo. Drawing on the Elaboration Likelihood Model (ELM), used text mining and the econometrics method to investigate the relationships among the central route (rebuttal's information readability and argument quality), peripheral route (rebuttal's source credibility, including authority and influence), and rebuttal acceptance, as well as the moderating effect of receiver's cognitive ability on these relationships. Our findings suggest that source authority had a negative effect on rebuttal acceptance, while source influence had a positive effect. Second, both information readability and argument quality had positive effects on rebuttal acceptance. In addition, individuals with low cognitive abilities relied more on source credibility and argument quality to accept rebuttals, while individuals with high cognitive abilities relied more on information readability. This study can provide decision support for practitioners to establish more effective fake news rebuttal strategies; it was especially valuable to reduce the negative impact of fake news related to major public health emergencies and safeguard the implementation of anti-epidemic strategies.

In 2022, Sedik *et al.* [7] have initiated a deep learning-based Fake News (FN) Detection method. This paper proposes a deep learning-based method for detecting FNs. The proposed system consists of three phases; text encoding, feature extraction, and classification. The text encoding process was carried out on the input news words using GLOVE for word representation. The encoded words are then embedded into a specific word length in order to be enrolled in the proposed deep learning models. The proposed deep learning models comprise both automatic feature extraction and classification tasks. Furthermore, this study proposes four different deep learning models, including CNNs and Concatenated CNNs (C-CNNs), LSTM, and GRU, to find an optimal model prior to the section of FNs that outperforms previous works. The proposed DL models are carried out on FNs and FNC datasets which are provided by kaggle, and the suggested C-CNNs algorithm obtained an accuracy and trained faster than others. Multiple evaluation metrics such as precision, recall, F1, and accuracy have been utilized to evaluate the outcome of the proposed models. The experimental results demonstrated overall improvements in the subject of FND when compared with the current models and validated the potential of the proposed methodology for the detection of FNs on Social Media (SM). This study will help researchers to broaden the knowledge of applications of CNNs based on DL methods for FND.

In 2021, Do *et al.* [8] have proposed a generic model that was able to take into account both the news content and the social context for the identification of fake news. Specifically, developed explore different aspects of the news content by using both shallow and deep representations. The shallow representations are produced with word2vec and doc2vec models while the deep representations were generated via transformer-based models. These representations are able to jointly or separately address four individual tasks, namely bias detection, clickbait detection, sentiment analysis, and toxicity detection. In addition, we make use of graph convolutional neural networks and mean-field layers in order to exploit the underlying structural information of the news articles. That way, developed were able to take into account the inherent correlation between the articles by leveraging their social context information. Experiments on widely-used benchmark datasets indicated the effectiveness of the proposed method.

**Problem Definition**

Due to the huge development arises in social media network, false news also get developed from classical texts. But, the user-created content published in social media platform is noisy, large-scale and multimodal. The existing approaches utilized for the fake news detection with deep learning approaches is showcased in Table 1. DAGA-NN [1] utilized only finite number of data for training to attain accurate detection rate in cross domain and it utilized feature extractor to attain effective basic feature from training dataset. But, it may get lagged when information are collected from different domain and minimize the effectiveness of system. Propagation2Vec [2] generated a new varying level for encoding propagation in the network and it easily detects the deadline in partial propagation approaches to perform effective reconstruction. Still, it didn’t have the capability to carry the domain relation with new labels and this leads to achieve poor performance rate. FND-SCTI [3] it has the ability to learn the shared multimodal representation and improve the textural information in new document. Yet, it became complex to classify the fake images from original images and it create complexity to the system. OPCNN-FAKE [4] provide enhanced performance rate in cross-validation and recognize the fake new effectively than existing approaches and it utilized glove word embedding to represent the features in matrix to enhance the detection rate. At the same time, it requires huge number of data to perform training and it didn’t encode the position of object. CNN and LSTM [5] effectively remove the features like noisy, redundant and irrelevant data from feature vector. But, it didn’t provide accurate results when large dataset is utilized. ELM [6] utilizes text mining in center route to get rebuttal acceptance from the receiver and achieve high information readability. Still, it didn’t have the capability to analyse the emotions presented in the text. CNN, LSTM and GRU [7] utilized minimal memory to store the data and also it achieved high accuracy than conventional approaches and trains the dataset effectively. But, it is more sensitivity when multiple random weights are used. GDMFN [8] has high effectiveness rate to smooth the node characters presented in the graph. Still, it requires more time to analysis the data and it delay the training process. So, it is essential to develop a new advanced system for fake news detection using deep learning approaches.

**Table 1:** Features and challenges of existing fake news detection model based on deep learning approaches

|  |  |  |  |
| --- | --- | --- | --- |
| **Author [citation]** | **Methodology** | **Features** | **Challenges** |
| Yuan *et al.* [1] | DAGA-NN | * It utilized only finite number of data for training to attain accurate detection rate in cross domain. * It utilized feature extractor to attain effective basic feature from training dataset. | * It may get lagged when information are collected from different domain and minimize the effectiveness of system. |
| Silva *et al.* [2] | Propagation2Vec | * It generated a new varying level for encoding propagation in the network. * It easily detects the deadline in partial propagation approaches to perform effective reconstruction. | * It didn’t have the capability to carry the domain relation with new labels and this leads to achieve poor performance rate. |
| Zeng *et al.* [3] | FND-SCTI | * It has the ability to learn the shared multimodal representation and improve the textural information in new document. | * It became complex to classify the fake images from original images and it create complexity to the system. |
| Saleh *et al.* [4] | OPCNN-FAKE | * It provide enhanced performance rate in cross-validation and recognize the fake new effectively than existing approaches. * It utilized glove word embedding to represent the features in matrix form to enhance the detection rate. | * It requires huge number of data to perform training and it didn’t encode the position of object. |
| Umer *et al.* [5] | CNN, LSTM | * It effectively removes the features like noisy, redundant and irrelevant data from feature vector. | * It didn’t provide accurate results when large dataset is utilized. |
| Wang *et al.* [6] | ELM | * It utilizes text mining in center route to get rebuttal acceptance from the receiver and achieve high information readability. | * It didn’t have the capability to analyse the emotions presented in the text. |
| Sedik *et al.* [7] | CNN, LSTM and GRU | * It utilized minimal memory to store the data. * It achieved high accuracy than conventional approaches and trains the dataset effectively. | * It is more sensitivity when multiple random weights are used. |
| Do *et al.* [8] | GDMFN | * It has high effectiveness rate to smooth the node characters presented in the graph. | * It requires more time to analysis the data and it delay the training process. |

**Research Methodology**

Fake news has recently leveraged the power and scale of online social media to effectively spread misinformation which not only erodes the trust of people on traditional presses and journalisms, but also manipulates the opinions and sentiments of the public. Society and individuals are negatively influenced both politically and socially by the widespread increase of fake news either way generated by humans or machines. In the era of social networks, the quick rotation of news makes it challenging to evaluate its reliability promptly. With the increasing popularity of social media, people have changed the way they access news. News online has become the major source of information for people. However, much information appearing on the Internet is dubious and even intended to mislead. Some fake news is so similar to the real ones that it is difficult for human to identify them. Therefore, automated fake news detection tools have become a crucial requirement. So, in this proposal, a novel deep learning approach will be developed for the recognition of fake news in pandemic situation. The major phases of the proposed model will be (a) Pre-processing, (b) Feature extraction, (c) optimal feature selection and (d) classification. Initially, text data will be collected from benchmark resources related to pandemic situation and provided to pre-processing phase. Then, the pre-processed data will be provided to feature extraction phase. The features will be extracted by using glove embedding, Bidirectional Encoder Representations from Transformers (BERT) and Term Frequency-Inverse Document Frequency (TFIDF). Later, the extracted features will be provided to fused optimal weighted feature selection phase, the weighted features will be selected and their weights will be optimized with improved Artificial Rabbits Optimization (ARO) [26]. The attained optimal weighted features will be subjected to classification phase. In classification phase, the fake news will be classified with the help of Optimal Multiscale Convolutional Neural-Long Short-Term Memory (OMCN-LSTM) and also their parameters in CNN and LSTM will be tuned by developed ARO. The developed (OMCN-LSTM) consist of convolutional layer, max pooling layer, LSTM layer and dense layer for processing the optimal weight feature to gain fake news classified output. Thus, the proposed model achieved effectively high accuracy rate in fake news detection than conventional approaches. The proposed fake news classification model is given in Figure.1.

Collected text

Feature extraction

Pre-processing

Fake news Classification

Fused optimal weighted feature selection

Glove embedding, BERT, TFIDF

T

Improved ARO

OMCN-LSTM

Classified outcome

Improved ARO

**Figure 1:** Proposed model of fake news classification

**Expected Outcome**

The proposed model will be implemented in Python and the results will be analyzed. Here, Type I measures are positive measures like Accuracy, Sensitivity, Specificity, Precision, Negative Predictive Value (NPV), F1Score and Mathews correlation coefficient (MCC), and Type II measures are negative measures like False positive rate (FPR), False negative rate (FNR), and False Discovery Rate (FDR).

**References**

1. RHua Yuan, Jie Zheng, Qiongwei Ye, Yu Qian and Yan Zhang, "Improving fake news detection with domain-adversarial and graph-attention neural network,"Decision Support Systems, Vol. 151, pp. 113633, December 2021.
2. Amila Silva, Yi Han, Ling Luo, Shanika Karunasekera and Christopher Leckie, "Propagation2Vec: Embedding partial propagation networks for explainable fake news early detection,"Information Processing & Management, vol. 58, no.5, pp. 102618, September 2021.
3. Jiangfeng Zeng , Yin Zhang and Xiao Ma, "Fake news detection for epidemic emergencies via deep correlations between text and images,"Sustainable Cities and Society, vol. 66, pp. 102652, March 2021.
4. H. Saleh, A. Alharbi and S. H. Alsamhi, "OPCNN-FAKE: Optimized Convolutional Neural Network for Fake News Detection," IEEE Access, vol. 9, pp. 129471-129489, 2021.
5. M. Umer, Z. Imtiaz, S. Ullah, A. Mehmood, G. S. Choi and B. -W. On, "Fake News Stance Detection Using Deep Learning Architecture (CNN-LSTM)," IEEE Access, vol. 8, pp. 156695-156706, 2020.
6. Xin Wang, Fan Chao, Guang Yu and Kaihang Zhang, "Factors influencing fake news rebuttal acceptance during the COVID-19 pandemic and the moderating effect of cognitive ability," Computers in Human Behavior, vol.130, pp. 107174, May 2022.
7. Ahmed Sedik, Amr A. Abohany, Karam M.Sallam, dKumudu Munasinghe and T.Medhat, "Deep fake news detection system based on concatenated and recurrent modalities," Expert Systems with Applications, vol. 208, pp. 117953, 1 December 2022.
8. T. Huu Do, M. Berneman, J. Patro, G. Bekoulis and N. Deligiannis, "Context-Aware Deep Markov Random Fields for Fake News Detection," IEEE Access, vol. 9, pp. 130042-130054, 2021.
9. P. Wei, F. Wu, Y. Sun, H. Zhou and X. -Y. Jing, "Modality and Event Adversarial Networks for Multi-Modal Fake News Detection," IEEE Signal Processing Letters, vol. 29, pp. 1382-1386, 2022.
10. K. A. Qureshi, R. A. S. Malick, M. Sabih and H. Cherifi, "Complex Network and Source Inspired COVID-19 Fake News Classification on Twitter," IEEE Access, vol. 9, pp. 139636-139656, 2021.
11. L. Ying, H. Yu, J. Wang, Y. Ji and S. Qian, "Fake News Detection via Multi-Modal Topic Memory Network," IEEE Access, vol. 9, pp. 132818-132829, 2021.
12. N. Seddari, A. Derhab, M. Belaoued, W. Halboob, J. Al-Muhtadi and A. Bouras, "A Hybrid Linguistic and Knowledge-Based Analysis Approach for Fake News Detection on Social Media," IEEE Access, vol. 10, pp. 62097-62109, 2022.
13. K. Xu, F. Wang, H. Wang and B. Yang, "Detecting fake news over online social media via domain reputations and content understanding," Tsinghua Science and Technology, vol. 25, no. 1, pp. 20-27, Feb. 2020.
14. T. Jiang, J. P. Li, A. U. Haq, A. Saboor and A. Ali, "A Novel Stacking Approach for Accurate Detection of Fake News," IEEE Access, vol. 9, pp. 22626-22639, 2021.
15. D. Li, H. Guo, Z. Wang and Z. Zheng, "Unsupervised Fake News Detection Based on Autoencoder," IEEE Access, vol. 9, pp. 29356-29365, 2021.
16. P. K. Verma, P. Agrawal, I. Amorim and R. Prodan, "WELFake: Word Embedding Over Linguistic Features for Fake News Detection," IEEE Transactions on Computational Social Systems, vol. 8, no. 4, pp. 881-893, Aug. 2021.
17. G. Shan, B. Zhao, J. R. Clavin, H. Zhang and S. Duan, "Poligraph: Intrusion-Tolerant and Distributed Fake News Detection System," IEEE Transactions on Information Forensics and Security, vol. 17, pp. 28-41, 2022.
18. B. Han, X. Han, H. Zhang, J. Li and X. Cao, "Fighting Fake News: Two Stream Network for Deepfake Detection via Learnable SRM," IEEE Transactions on Biometrics, Behavior, and Identity Science, vol. 3, no. 3, pp. 320-331, July 2021.
19. X. Dong, U. Victor and L. Qian, "Two-Path Deep Semisupervised Learning for Timely Fake News Detection," IEEE Transactions on Computational Social Systems, vol. 7, no. 6, pp. 1386-1398, Dec. 2020.
20. K. A. Qureshi, R. A. S. Malick, M. Sabih and H. Cherifi, "Complex Network and Source Inspired COVID-19 Fake News Classification on Twitter," IEEE Access, vol. 9, pp. 139636-139656, 2021.
21. L. Ying, H. Yu, J. Wang, Y. Ji and S. Qian, "Multi-Level Multi-Modal Cross-Attention Network for Fake News Detection," IEEE Access, vol. 9, pp. 132363-132373, 2021.
22. P. Li, X. Sun, H. Yu, Y. Tian, F. Yao and G. Xu, "Entity-Oriented Multi-Modal Alignment and Fusion Network for Fake News Detection," IEEE Transactions on Multimedia, vol. 24, pp. 3455-3468, 2022.
23. C. Li et al., "Joint Stance and Rumor Detection in Hierarchical Heterogeneous Graph," IEEE Transactions on Neural Networks and Learning Systems, vol. 33, no. 6, pp. 2530-2542, June 2022.
24. G. Shrivastava, P. Kumar, R. P. Ojha, P. K. Srivastava, S. Mohan and G. Srivastava, "Defensive Modeling of Fake News Through Online Social Networks," IEEE Transactions on Computational Social Systems, vol. 7, no. 5, pp. 1159-1167, Oct. 2020.
25. N. R. de Oliveira, D. S. V. Medeiros and D. M. F. Mattos, "A Sensitive Stylistic Approach to Identify Fake News on Social Networking," IEEE Signal Processing Letters, vol. 27, pp. 1250-1254, 2020.
26. Liying Wang, Qingjiao Cao, Zhenxing Zhang, Seyedali Mirjalili and Weiguo Zhao, "Artificial rabbits optimization: A new bio-inspired meta-heuristic algorithm for solving engineering optimization problems," Engineering Applications of Artificial Intelligence, vol.114, pp. 105082, September 2022.