





Data & Knowledge Engineering

Volume 144, March 2023, 102124

Meta-heuristic Searched-Ensemble Learning for fake news detection with optimal weighted feature selection approach

S. Hannah Nithya (Research Scholar)  , Dr. Arun Sahayadhas (Professor)

Show more 

 Share  Cite

<https://doi.org/10.1016/j.datak.2022.102124> 

[Get rights and content](#) 

Abstract

Nowadays, fake news has turned into a major problem because of the negative impact that it creates on society. Social media allows people to spread information on the internet with slight investigations and to add fewer filters than the actual content. Nowadays, the false news in the internet community is unsure, and it creates a wrong impression among the users. Detecting false news has become a critical task based on shared content. To tackle the false news growth in social media, various automatic detection schemes were evaluated. The “Natural Language Processing (NLP)” method also gives a prominent solution for false news detection. The main intention of this paper is to design and introduce an innovative false news recognition method using Meta-heuristic Searched-Ensemble Learning (MS-EL). Further, the selected features are extracted by the “Term Frequency-Inverse Document Frequency (TF-IDF)” and also Word2vec features. Here, the extracted selected features are integrated with the Hybrid Squirrel–Dragonfly Search Optimization (HS-DSO) is used to optimize the weighted feature selection approach with the fitness function of solving data variance and correlation. The proposed MS-EL is adopted in the classification part, having

three sets of classifiers, Long Short-Term Memory (LSTM), Support Vector Machine (SVM), and Deep Neural Network (DNN). Here, the ensemble classifier is enhanced by the same HS-DSO that shows the parameter tuning with a high convergence rate. From the experimental outcomes, the accuracy of HS-DSO-MS-EL is 22% higher than BMO-MS-EL, 24% higher than SP-BMO-MS-EL, 30% higher than SSA-MS-EL, and 29% higher than DA-MS-EL. Thus, the experimental analysis with standard datasets establishes that the introduced fake news detection method has gained higher accuracy than the existing models.

Introduction

In recent times, the capability of network users to write everything on social media and website articles has led people to mislead the information[1]. The internet community platform, such as youtube, Facebook, Twitter, Instagram, etc., has become the main source of information for people worldwide, especially in developing countries[2]. Through the spread of false news, the whole society is damaged in various sources like business, culture, and politics, the fake news is beneficial or damaging at the same time[3]. Social media users not only cause consequences but also help people to connect through a network and offer more prospects regarding reputation and career opportunities, either directly or indirectly[4]. Alternately, social media produces a major platform for the formation and dissemination of false news. False news is highly influential and can spread extremely fast[5]. Day by day, the explosion of new information is increased with the utilization of internet community users. The false information is tricky to correct and has long-lasting impacts. People expose their information on social media, and if the information is not accurate, they expose them with lies. The misinformation spread widely can cause damage to large corporations, stock markets, businesses, etc., and harm the people[6]. The detection of fake news has been essential in internet community websites to avoid the spreading of fake news.

Fake news is usually denoted as yellow press, which intentionally spreads misinformation through traditional printed newspapers and online social media[7]. The inherent attributes of the internet community, like easy access, less cost, and fast propagation, can mislead the judgment of the public, upset society, and can also spoil the credibility of social media[8]. So, it is essential to investigate false news recognition. Recognizing false news from media and identifying the creators and the content must be more accurate to eradicate the large source of false news from the origin of the internet community[9]. The traditional fake news detection algorithm poses unique attributes and limitations, which makes the existing news media irrelevant and ineffective. The motive of false news is to make the readers believe the misinformation, where the content of fake news is difficult to detect, and it is

essential to consist of auxiliary information. The auxiliary information is challenging because the users produce data that are huge, noisy, and unstructured[10]. Detecting false news in the internet community is challenging, and therefore, numerous deep structured learning and knowledge engineering approaches were incorporated to identify the origin of fake news[11].

In recent years, various techniques have been proposed to identify fake news in online media, including conventional learning approaches[12]. The existing techniques like “Random Forest (RF), Decision Tree (DT), and Support Vector Machine (SVM)” are used to recognize false news, which is time-consuming and labor-demanding[13]. For example, SVM classifies the fake post and article and utilizes time series composition to develop characteristic social variations[14]. The neural network has developed a huge success in fake news detection; therefore, traditional deep learning models have attained better performance than other heuristic algorithms with high stability[15]. The fake post is extracted and identified with the existing works. The Recurrent Neural Networks (RNN) and LSTM are employed to extricate the language characters for false news detection[16]. The high level characteristics from the news and posts are extracted to identify social media platform sources. However, the improved models contain noisy information and other disturbances in content[17]. Hence, ME-SL based false news detection approach is introduced in this proposed paper.

The major contribution of the suggested method of detecting false news is planned here.

- To design fake news detection model using ME-SL with optimized feature selection using the proposed method of HS-DSO. The HS-DSO is utilized to optimize the fitness solution for solving the data.
- To extract the features from TF-IDF and word2vec features is to develop a novel weighted feature selection by adopting HS-DSO with the purpose of reducing the correlation and maximizing the variance. Here, the weight is analyzed with the help of TF-IDF in the feature extraction and helps retrieve the information.
- To suggest a hybrid HS-DSO algorithm for optimizing the weighted features and enhancing accurate fake news detection. The proposed method helps to attain the optimal solution with a high convergence rate.
- To identify the detection of the real and fake news with ME-SL by optimizing the hidden neuron count of DNN and NN, several epochs of SVM, the learning

rate of LSTM, and several epochs of ensemble using the HS-DSO algorithm to increase the accuracy of detection.

The balance sections of the suggested method of detecting fake news are specified here. The literature review is briefly discussed in Section 2. Section 3 specifies the fake news dataset description and preliminary processing steps for detection. The architectural model and optimal weighted feature selection using HS-DSO are discussed in Section 4. Section 5 specifies the MS-EL for enhanced fake news detection. The results and discussion are explained in Section 6, and the conclusion of the proposed work is discussed in Section 7.

Access through your organization

Check access to the full text by signing in through your organization.

Access through **your organization**

Section snippets

Related works

Existing deep learning-based fake news detection methods

In 2021, Saleh et al. [18] has developed a novel machine knowledge learning and deep structured learning-based method for addressing the false news on online media platforms. The major intention of this paper was to attain optimal solutions and high accuracy performance. Here, to detect fake news, a novel Optimized Convolutional Neural Network (OCNN) was introduced. The suggested method was gathered from other existing learning methods....

Fake news dataset illustration

The benchmark datasets for suggesting the input data are gathered from KaiDMML-politifact, KaiDMML-gossipcop, and the Stance detection dataset for FNC-1.

Dataset 1: The standard dataset is gathered from the website “<https://github.com/KaiDMML/FakeNewsNet>”-Access ↗ date: 9-03-2022. The tool which is collected from the fake news dataset is the FakeNewsTracker. This tool is essential for

collecting the data, analyzing, and visualizing fake news. Here, the fake news is downloaded from the websites,...

Proposed fake news detection architecture

False news has attained great attraction from researchers, and the recognition of false news is required with better efficiency. Fake news is widely extended from the source of multi-media in the formation of content and news context through various user-based features. Different techniques are implemented for recognizing the false news collected on social media platforms, and also false news detection [29] has been a major challenge in all corners of the world. The identification of false news ...

Classifiers used for ensemble learning

The classifiers used for proposed ensemble learning are explained here.

SVM[36]: It is applied to various classification and detection applications and has the largest margin to separate the hyperplane samples into high-dimension space. Estimate the dataset $U = \{(w_{o1}, x_{o1}), (w_{o2}, x_{o2}), \dots, (w_{oS_F}, x_{oS_F})\}$, where the terms $z_{o_{io}} \in \mathcal{R}^{p_o}$ $y_{o_{io}} \in \mathcal{R}^{p_o}$ were used to divide the hyperplane $z_o^t x_o + y_o = 0$, in which the direction of the hyperplane is given z_o , and the distance between the origin and hyperplane is denoted as y_o

Simulation analysis

The introduced detection of false news was accomplished in “MATLAB 2020a”, and the simulation analysis was determined. The population size was ten, and the highest iteration was 25. The introduced model was evaluated over other conventional algorithms like Barnacles Mating Optimizer (BMO)[39], SP-BMO[37], SSA[32], and DA[35] and further evaluated with other existing learning models like DNN[36], SVM[38], LSTM[27], LSTM-DNN-SVM[27], [36], [38] and SP-BMO-LSTM[37] concerning the positive ...

Conclusion

The novelty of this research work is to detect and identify the fake news detection model developed with a novel MS-EL enabled with an optimal weighted feature selection method. The congregated input dataset was preprocessed with stop word removal, blank space removal, and stemming to remove the unwanted data and amplify the noises. Further, the

pre-processed data were subjected to a feature extraction task, which was performed with TF-IDF and word2vec. Later, the features extracted were given...

CRedit authorship contribution statement

S. Hannah Nithya: Design and implementation of the research, Analysis of the results, Writing of the manuscript. **Dr.Arun Sahayadhas:** Design and implementation of the research, Analysis of the results, Writing of the manuscript....

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper....

Acknowledgments

I wish to thank my parents for their support and encouragement throughout my study....

Funding statement

None....

Ethics statement

This paper does not contain any studies with human participants or animals performed by any of the authors....

Hannah Nithya

I Hannah Nithya works for a 24/7 news channel called Sathiyam television as a news presenter and a reporter. I have a M.tech degree from Karunya University. I am currently pursuing my doctorate degree in Vels University under the esteem guide ship of Dr. Arun Sahayadhas....

[Recommended articles](#)

References (39)

JangS.M. *et al.*

Third person effects of fake news: fake news regulation and media literacy interventions

Comput. Hum. Behav. (2018)

ZhangChaowei *et al.*

Detecting fake news for reducing misinformation risks using analytics approaches

European J. Oper. Res. (2019)

DavoudiMansour *et al.*

DSS: A hybrid deep model for fake news detection using propagation tree and stance network

Expert Syst. Appl. (2022)

LiuJinshuo *et al.*

DTN: Deep triple network for topic specific fake news detection

J. Web Semant. (2021)

SongChenguang *et al.*

A multimodal fake news detection model based on crossmodal attention residual and multichannel convolutional neural networks

Inf. Process. Manage. (2021)

ChoiHyewon *et al.*

Effective fake news video detection using domain knowledge and multimodal data fusion on youtube

Pattern Recognit. Lett. (2022)

BerghelH.

Alt-news and post-truths in the fake news era

Computer (2017)

ChangB. *et al.*

Study on information diffusion analysis in social networks and its applications

Int. J. Autom. Comput. (2018)

ShuK. *et al.*

Fake news detection on social media: A data mining perspective

SIGKDD Explor. Newslett. (2017)

VosoughiS. *et al.*

The spread of true and false news online

Science (2018)



[View more references](#)

Cited by (10)

[What makes sentiment signals work? Sentiment and stance multi-task learning for fake news detection](#)

2024, Knowledge-Based Systems

[Show abstract](#)

[A comprehensive survey of fake news in social networks: Attributes, features, and detection approaches](#)

2023, Journal of King Saud University - Computer and Information Sciences

[Show abstract](#)

[ChatGPT and ensemble learning for mobile text misinformation identification](#) ↗

2024, Emerging Technologies for Securing the Cloud and IoT

[Fake news detection using ensemble techniques](#) ↗

2024, Multimedia Tools and Applications

[Misinformation Detection Using an Ensemble Method with Emphasis on Sentiment and Emotional Analyses](#) ↗

2023, Proceedings - 2023 IEEE/ACIS 21st International Conference on Software Engineering Research, Management and Applications, SERA 2023

[Regulating Fake News and Hoaxes: A Comparative Analysis of Indonesia and Malaysia](#) ↗

2023, Journal of Human Rights, Culture and Legal System



[View all citing articles on Scopus ↗](#)

Hannah Nithya

I Hannah Nithya works for a 24/7 news channel called Sathiyam television as a news presenter and a reporter. I have a M.tech degree from Karunya University. I am currently pursuing my doctorate degree in Vels University under the esteemed guidance of Dr. Arun Sahayadhas.

[View full text](#)

© 2022 Elsevier B.V. All rights reserved.



All content on this site: Copyright © 2024 Elsevier B.V., its licensors, and contributors. All rights are reserved, including those for text and data mining, AI training, and similar technologies. For all open access content, the Creative Commons licensing terms apply.

