



# A framework of fake news detection on web platform using ConvNet

Dinesh Kumar Vishwakarma<sup>1</sup> · Priyanka Meel<sup>1</sup> · Ashima Yadav<sup>2</sup> · Kuldeep Singh<sup>3</sup>

Received: 29 November 2022 / Revised: 28 December 2022 / Accepted: 7 January 2023 / Published online: 16 January 2023  
© The Author(s), under exclusive licence to Springer-Verlag GmbH Austria, part of Springer Nature 2023

## Abstract

Social media and web have become popular platforms of information sharing, knowledge gathering, expressing sentiments, perceiving choices regarding products and services through major news sources, and an active channel for marketing. Hence, with these promising features comes the threat of misinformation propagation, leading to undesirable effects. Therefore, accurate verification of fraudulent content on time is in high demand. Hence, to tackle this problem, we proposed a novel framework WSCH-CNN (web scrapping content heading CNN) model which counters the issue of fake (or mislead) news using convolutional neural networks (CNN). The framework consists of two CNN models named content model and heading model, which are used to find the linguistic similarities in fake news, and classifies them into real news or fake news. Both the models are evaluated on two publicly available datasets, namely Kaggle dataset and fake news challenge dataset, and two self-compiled real-world datasets, namely text dataset (text dataset of news articles) and multimedia dataset (Image dataset compiled from Facebook and Twitter), using accuracy, precision, recall, and F1 score as evaluation metrics. Moreover, the recognition accuracy achieved on these datasets is compared with similar state-of-the-art results. The proposed WSCH-CNN model proved quite successful in detecting the fake news with a high accuracy of 85.06% on multimedia dataset, 94.16% for heading model and 85.32% for content model which supersedes the state-of-the-art.

**Keywords** Authenticity of news images · Convolutional neural network · Classification of the news · Fake news · Social media

## 1 Introduction

Recent days, it has been observed that social media platforms are becoming a powerful tool for dissemination of information to the users. In present days, social media is

one of the most emerging media platforms and widely used throughout the world (Nisar et al. 2019). Online platforms provide freedom of expression to the users, which gives rise to a massive carnival of user-generated data present in the form of text, image, audio, and video (Viviani and Pasi 2017). These unverified and unauthenticated volumes of information are creating an epidemic of fake news, which led to the problem of users being misled. The total effect of being misled on social media has a very adverse effect in terms of sentiments of people, their perceptions, and choices being made in the future (Wessel et al. 2016). Therefore, social media platforms need to reach an equilibrium state in the tradeoff between an individual's freedom to post and welfare of the public.

The proliferation and real-time effect of social media are such that breaking news appears on microblogs first, before making it through the traditional media outlet. The news related to the death of Osama bin Laden was first broke out on Twitter, and millions had rapidly dispersed it through their Twitter and Facebook pages long before the US president officially announced it on the public media. Along

✉ Dinesh Kumar Vishwakarma  
dvishwakarma@gmail.com

Priyanka Meel  
priyankameel86@gmail.com

Ashima Yadav  
ashimayadavdtu@gmail.com

Kuldeep Singh  
kuldeep.ece@mnit.ac.in

<sup>1</sup> Department of Information Technology, Delhi Technological University, New Delhi, India

<sup>2</sup> School of Computer Science Engineering and Technology, Bennett University, Greater Noida, India

<sup>3</sup> Department of Electronics and Communication Engineering, Malaviya National Institute of Technology, Jaipur, Rajasthan, India



**Fig. 1** Images of fake news on social media

with the text, images are also circulated on microblogging platforms to strengthen the authenticity of a breaking news story, as some platforms support fixed-length text messages. Debunking fake news, both in text and multimedia format at an early stage of diffusion, is particularly crucial to minimize their harmful effects (Singh and Sharma 2022; Garg and Sharma 2020).

Figure 1 shows some famous examples of fake news articles which are embedded in images (courtesy Facebook). Every day in our life, each one of us encounters fake news in one way or the other, which affects us in some way. It was observed that during the US presidential election, there were thousands of the times the fake news were circulated over social media platforms on the US presidential election. Tang et al. (2011) and Cheung et al. (2015) demonstrations how content circulates over the social media, and the fake news is a chunk of such content. What makes the situation more complicated is “name theft of authoritative media” (Niu 2008), which has more effect on content users as it makes them believe that the event is real. Many times, it has been seen that the trending topics (Aiello et al. 2013) on social media are reported as fake news.

The problem of fake news is now getting the attention of many researchers (Meel and Vishwakarma 2019), and the researchers are working hard to come up with a solution. The study in this field can be broadly divided into two categories, such as algorithms used for processing text data in the image and algorithms used for processing only the text data. The image-based algorithms extract various features of the images to train a model to classify images based on these features. The text-based algorithm mainly uses text pattern and matches them with already existing patterns of fake news, which is popularly termed as Linguistic methodology (Sharma et al. 2022; Singh and Sharma 2021).

The earlier works in the field of fake news mainly focused on processing the input data in the form of images and extracts visual and statistical features (Jin et al. 2017) from the image related to the news. Elkasrawi et al. (2016) process the previous occurrences of the image on the web to check its authenticity. The question which remains unanswered is given as follows: what if there is no image associated with the news event? In such cases, present system fails to provide an accurate classification. To overcome this limitation, we propose a model to tackle the problem of fake news with textual data, which can further be used with images.

## 2 Research objectives

The main goal of this work is to develop a novel framework which helps in truth analysis of information available on the social media platforms in the form of images and text. In recent time, CNN-based text classification has taken its toll and is being proved quite useful in text classification and categorization (Kim 2014; Conneau et al. 2016; Jhonson and Zhang 2015). The simplicity and high prediction accuracy of CNN make it suitable for handling the problem of detection of fake news. To our best knowledge, this is the first attempt to analyze the problem of fake news for multimedia data in which instead of using statistical and visual image features, image text is being extracted, and web scraping technique is used with verification of links by developing a verified sources model. Finally, headlines and text from the verified links are being analyzed with the CNN model separately. The proposed framework consists of two CNN-based models, namely content model (CM) and heading model (HM). A model is designed and developed after a feasibility study of ground truth, which includes earlier state-of-the-arts. The

significant contributions of the proposed work are summarized as follows:

- We propose a novel framework WSCH-CNN (Web Scrapping Content Heading CNN) model for detecting the fake news which spreads through multimedia data on social media platforms like Facebook and Twitter. The entire framework is divided into four major units: processing unit, verified sources model, CNN model, and decision-making unit. We have developed deep learning-based CNN models, namely content model (CM) and heading model (HM) for separately classifying the headings and contents of the reliable links into fake and real. Both the CNN-based models are trained on Word2Vec model.
- We have manually prepared a bag of word model called as verified sources model (VSM), which contains a list of authenticated media houses and newspapers. The VSM model categorizes a link into reliable or unreliable.
- In order to test the efficacy of the proposed WSCH-CNN framework, the model is evaluated on two publicly available text-based datasets, namely Kaggle dataset (KD) and fake news challenge dataset (FNCD). Both the CM and HM are evaluated on these datasets based on evaluation metrics like accuracy, precision, F1 score, and recall.
- Moreover, to test the robustness of the proposed framework, we have scrapped the web to create two real-world datasets: text dataset (TD) which contains 12,689 news articles, out of which 5689 are fake news, and 7000 are real news, and multimedia dataset (MD) which contains 8496 real news event images and 7168 fake news event image. The proposed framework is also evaluated on these datasets.
- To demonstrate the effectiveness of our approach, we have also performed a comparison study with other state-of-the-arts. The experimental results suggest that there exists a linguistic pattern between fake news as fake news writers opt a particular style of writing to persuade readers to believe them as genuine.

The rest of the paper is structured as follows: Sect. 2 describes the literature survey and work done by other researchers in this field. The proposed model is explained in Sect. 3. The experimental work and discussion of results are presented in Sect. 4. Lastly, Sect. 5 concludes the work.

### 3 Related works

This section explores the related work done by various researchers in the field of fake news detection. Currently, researchers have employed many approaches (Jang et al. 2018) to detect real news content. Figueira and Oliveira

(Figueira and Oliveira 2017) proposed an algorithm which tries to identify facts and validate them to ensure sources are reliable. However, some challenges are to find the best infrastructure. An unsupervised causality-based framework is introduced in Shaabani et al. (2018) that also does label propagation. This approach identifies the pathogenic user accounts, which spread fake news without using cascade path information, network structure, content, and user's information. This method can be combined with complementary supervised techniques for enhanced results. Analysis of implicit and explicit profile features between the user groups is performed in Shu et al. (2018), which reveals their potential to differentiate between real news and fake news. These features are aggregated to detect fake news. Alrubaian et al. (Alrubaian et al. 2018) proposed a system that consists of four components: a reputation-based component, a user experience component, a credibility classifier engine, and a feature-ranking algorithm, which operates together in an algorithmic form to assess the credibility of Twitter content. Buntain and Golbeck (2017a) developed a classification model to predict whether a thread of Twitter conversation is accurate or not, using the features inspired from existing work on the credibility of Twitter stories. However, the model is dependent on structurally different datasets and popular Twitter threads. Existing social technologies solutions for reducing the creation and spreading of fake news is discussed in Campan et al. (2017) and authors in Egele et al. (2017) propose a technique to identify compromises of high-profile accounts. High-profile accounts usually show consistent behavior over time. It can detect and prevent real-world attacks against popular companies. However, there is a possibility that an attacker who is aware of all these techniques can prevent compromised accounts from detection. The datasets on machine learning algorithms like support vector machines, gradient boosting, stochastic gradient descent, and random forests (Gilda 2017b) are being tested. Stochastic gradient descent model identified non-credible sources with an accuracy of 77.2%. (Aiello et al. 2013; Jin et al. 2017) analyzed the critical role of image content in automatic news verification on microblogs. The 7% higher accuracy has been achieved compared with baseline approaches using only non-image features. These methods can be combined with existing non-image features to be able to detect media-rich news. The authors in Granik and Mesyura (2017) applied naive Bayes classifier for fake news detection. However, this can be improved by getting more data, using the dataset with lengthy news articles, removing stop words from articles, using stemming, and groups of words instead of separate words for calculating probabilities.

Recent studies are more focused on deep learning-based techniques (Singh and Sharma 2021a; b). Long short-term memory (LSTM) networks with pooling operation of CNN (Liu et al. 2019) are being used for building the rumor

identification model by taking into account the diffusion structure, forwarding contents, and spreaders. To investigate the performance of the model, they conducted experiments in two modes: *Afterward Identification Mode* and *Halfway Identification Mode*. Experimental results on Sina Weibo<sup>1</sup> dataset suggests that the proposed model can learn contextual and hidden information of diffusion structure, forwarding contents, and spreaders. Researchers in Barbado et al. (2019) proposed a new feature framework for identifying the fake reviews in electronics domain by distinguishing review centric features and user-centric features. User-centric features focus on the information regarding how users behave in a particular social network and which type of data they provide. These features are divided into four types: personal features (P), social features (S), reviewing activity features (RA) and trusting features (T). Review centric features use textual metrics such as average lengths, average words per sentence, term frequency-inverse document frequency (TF-IDF), word2vec, emotion analysis, and sentiment analysis to extract the lexical and syntactic features.

The results prove that user-centric features can give better F-scores to detect fake reviews. The area of fake news and rumor detection is gaining a lot of popularity among researchers. Many authors have presented a comprehensive survey in this area. Several reviews on online fake news detection exist in the literature. (Liu et al. 2019) gives a complete overview about state-of-the-art methods, discussing the feature engineering process by enlightening the user-specific, content-specific, and context-specific features, and provides a detailed study about the existing datasets that are applied for classifying the fake news. Similarly, (Bondielli and Marcelloni 2019; Sharma et al. 2019; Zubiaga et al. 2018) provides an in-depth survey which summarizes all the work done in the area of fake news and rumor detection. Many of the previous works have assumed rumors to be false. However, Alkhodair et al. (2019) demonstrated that rumors might be false at the time of detection, though they may be considered true later with time. They developed a model which can detect long-lasting rumors that spread in the social media by training word embeddings and passing them to the joint LSTM-RNN model. Experimental results on real-world datasets show that the proposed method outperformed the state-of-the-art works. A novel framework which could predict the future topics which are vulnerable to misinformation and could result in fake news topics is introduced in Vicario et al. (2019). They applied the following series of features on the Facebook post: presentation distance, mean response distance, controversy of an entity, perception of an entity, captivation of the entity and finally

classified the entity as covered or uncovered for generating the probable targets for fake news.

## 4 Proposed model

The following section explains the modules present in the proposed model. The proposed model is composed mainly of four units: (1) processing unit, (2) verified sources model (VSM), (3) CNN models, and (4) decision-making unit (DMU). The processing unit performs three micro-tasks, namely extracting text from the image, entity extraction, and scraping the Google results. The verified sources model is a self-compiled list of authentic media houses and newspapers. The CNN models are responsible for classifying the heading and content of a news article, and the decision-making unit is responsible for making the final determination, after taking inputs from two CNN models. The objective of the anticipated model is to authenticate the news events that are trending on social media platforms. Figure 2 shows the flow diagram of the proposed WSCH-CNN( Web Scrapping Content Heading CNN) model. In the following section, the details about each module are discussed in brief.

### 4.1 Processing unit

The first module is “Processing unit”, and key words are the extraction of the text from the images, followed by extraction of entities from the extracted text, and searching the entities on search engine and finally collecting the results. The text extraction is done via (Yih et al. 2016). The key steps involved in the algorithm employed by Yih et al. (2016) are the detection, enhancement, stroke width, and filtering. An optical character recognition (OCR) is used to extract the text from the detected region.

For entity extraction, defined algorithms and grammar are used. A text cleaning is used before extracting the entries. The process of text cleaning includes removal of all non-alphabetic characters, eliminating multiple occurrences of the words, checking the validity of English word, checking each word for spelling errors with 1-edit distance, and elimination of media houses name or newspaper name to eliminate bias. Extracted entities formulate a query, which is searched on Google, and results are scrapped for the same.

### 4.2 Verified sources model (VSM)

Social media provides an open platform where anybody can express his/her sentiments, and emotions without any verification on its genuineness, and there are a larger number of content producers, who produce content without any fact check. However, there are many trustworthy media houses whose honesty cannot be questioned and the content

<sup>1</sup> <https://www.weibo.com/overseas>.

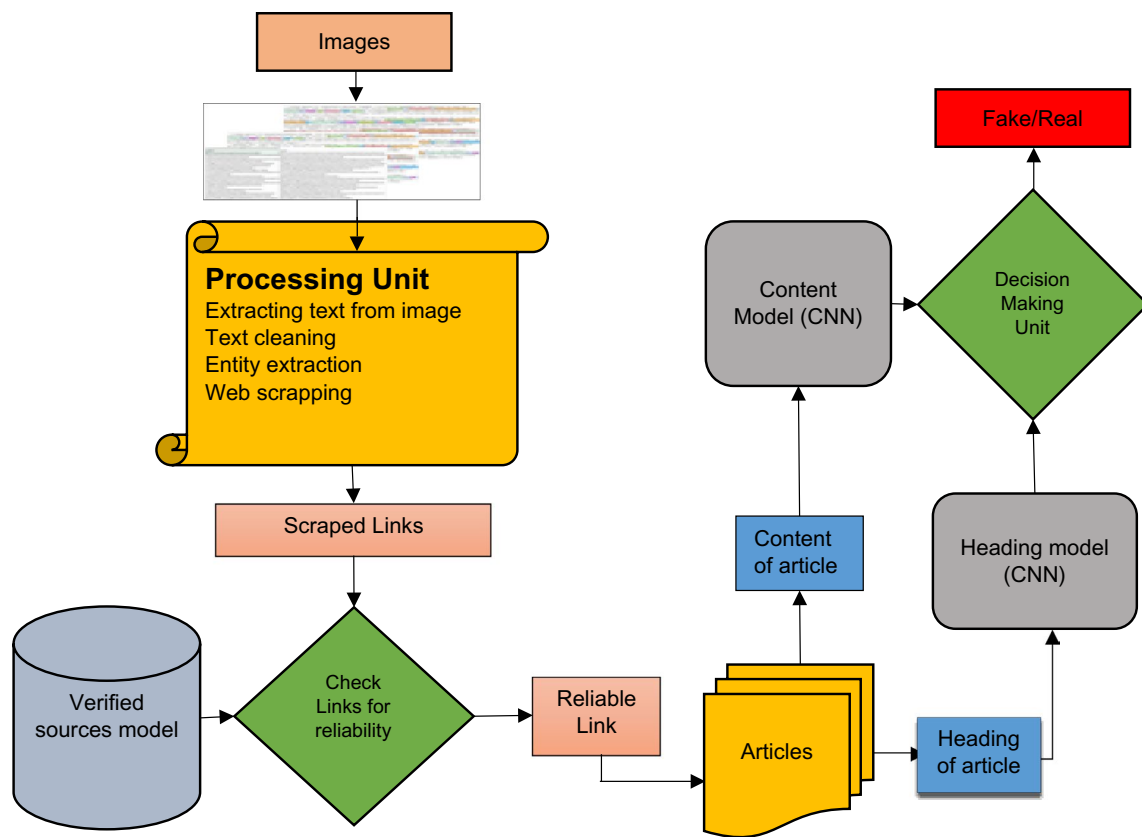


Fig. 2 Proposed WSCH-CNN (web scrapping content heading CNN) model architecture

produced by them is done after verification of the facts. Considering the significance of their news, a double-check may be mandatory. The media is considered to be the fourth pillar of democracy. Hence, the news channel like Fox News, CNN, NDTV, AAJTAK, etc. and newspaper websites like The Hindu, the Times of India, The Washington Post, New York Daily News, etc. serves as the credible sources for our model. A list of several such newspapers and media houses has been compiled, and it is called verified sources model (VSM). Hence, it is the bag of a manually gathered list of all the authenticated media houses and newspapers which is used to classify a link into reliable and unreliable. The list consists of entries of the international, national, and local newspapers and media houses.

### 4.3 CNN models

Convolutional neural network (Conneau et al. 2016) started their journey by their application in the field of image processing. They have proved very useful in the field of NLP, and an example of that is seen in Yih et al. (2014) where it shows the tremendous result in semantic parsing and search query retrieval (Shen et al. 2014). Both CNN models have been trained on manually collected text data of fake and

real news. Our CNN models are adapted from Kim (2014). The hyperparameters used for our CNN model are given as follows:

- To take into consideration the elements of a matrix which do not have adjacent elements to their top and left, we have used zero-padding. Adding zero-padding is also called *wide convolution*.
- A key aspect of convolutional neural networks is *pooling layers*, typically applied after the convolutional layers. We used max-pooling over the size of the stride of one step.
- L2 norm constraints on the weight vectors have not been enforced. Other parameters like filter window, dropout used are the same as in Kim (2014).

Both the content model (CM) and headings model (HM) use the same hyperparameters. CM is trained to classify the content of the article into two categories: i) fake and ii) real. HM is trained to classify the title of the article into two types: i) positive (if the title is telling that the news is a hoax or fake), and ii) negative. The training data were preprocessed to suit the need of the algorithm. For training, pre-trained word2vec (Jhonson and Zhang 2015) is used, which is available for

free. Max pooling is applied to downsize the input matrix, i.e., reduce the dimensions of the input. Max pooling is applied by using a fixed size filter, which traverses over the matrix with a fixed window size, known as Stride. A filter of size  $2 \times 2$  is iterating over the matrix and is taking the max value of all value that comes within the range of the filter. A  $2 \times 3$  matrix is downsized to  $1 \times 2$ .

#### 4.4 Decision-making unit

This module takes the final decision based upon the outputs of the content model and headings model. If any of the reliable links says, that the event is a hoax or fake, it is classified as fake, else Eq. 1 is used to find the reality parameter (Rp). Reality parameter is defined as the ratio of links, whose content is classified as real (Cr) by CNN models to that of total search results (Tsr) taken (i.e., Google search results). By manual inspecting a large number of queries (and their search results), it has been found that top 15 results provide an accurate overview of the query. If Rp is greater than 0.5, then the news is classified as real.

$$R_p = \frac{C_r}{T_{sr}} \quad (1)$$

where Rp=Realty parameter, Cr=Links classified as real by content and headings models, Tsr=Total search results.

Algorithm 1 describes the proposed system and how CNN models are used to classify images. The *ExtractTextFromImage* extracts the text from an image, which goes as an input to *EntityExtractor*. The *EntityExtractor* is responsible for extracting entities from text. These extracted entities are searched on Google, and a list of top 15 results are returned by the *GetGoogleSearchResults*. This obtained list goes through our bag-of-words (VSM) to classify them as reliable and unreliable, where the list of unreliable links are discarded. Further, the reliable links are iterated one by one for classifying their heading and content. After the content and heading of an article have passed through our model, the reality parameter is calculated, whose value if above the threshold, the event is classified as real.

---

#### Algorithm 1

---

```

1: text = ExtractTextFromImage (image)
2: query = EntityExtraction(text)
3: results = GetGoogleSearchResults(query)
4: reliableLinks = VerifiedSourcesModel(results) [VSM]
5: numberOfLinks = length(reliableLinks)
6: counter= 0
7: Foreach link in reliableLink:
8:     content, heading = ScrapeTheLink(link)
9:     cm = ContentModel(content)
10:    hm = HeadingModel(heading)
11:    if hm is true:
12:        return "Classified as Fake news"
13:    if cm is real:
14:        counter += 1
15:    reality parameter = counter/ len(results)
16:    if (ratio) > 0.5)
17:        return "Classified as Real news"

```

---

## 5 Experimental work and results

To evaluate the performance of the proposed model, we conduct an experiment on two publically available datasets and two self-created datasets. The publically available datasets are Kaggle's fake news dataset (Kaggle 2017) and fake news data (Fakenewschallenge 2017). The performance of the algorithm is measured in terms of detection accuracy,

precision, recall, F1-score, and calculated using Eqs. 2–5, respectively.

$$\text{Detection accuracy} = \frac{TN + TP}{TN + TP + FN + FP} * (100) \quad (2)$$

$$\text{Precision} = \frac{TP}{TP + FP} * (100) \quad (3)$$

**Table 1** Dataset details

Dataset name	Details	Total entries	Fake news	Real news
Kaggle dataset (KD)	Text dataset consisting of news articles with their content and heading	14,565	6850	7715
Fake news challenge dataset (FNCD)	Contains pair of heading and content for each news article	28,861	17,168	11,693
Text dataset (TD)	Self-compiled text dataset of news articles scraped from the web, contains heading and content of the article	12,689	5689	7000
Multimedia dataset (MD)	Multimedia (images) test dataset, compiled from microblogs sites (Facebook and Twitter)	15,664 images	7168	8496

A	B	C
Headline	Content	Label
Cast Your Vote: Whose Foreign Policy Position Do You Prefer?	Cast Your Vote: Whose Foreign Policy Position Do You Prefer? Posted on Oct 28, 2016 Donald Trump and Hillary Clinton	TRUE
Robert Scheer and Amy Trask Discuss Gender, Race and Profit in the World of Professional Sports	Robert Scheer and Amy Trask Discuss Gender, Race and Profit in the World of Professional Sports Posted on Oct 28, 2016	TRUE
Facebook Lets Advertisers Exclude Users by Race	Facebook Lets Advertisers Exclude Users by Race Posted on Oct 29, 2016	TRUE
Photo of the Week: President Who Let Citigroup Staff His Cabinet Calls for Solidarity From Voters	Photo of the Week: President Who Let Citigroup Staff His Cabinet Calls for Solidarity From Voters Posted on Oct 29,	TRUE
Electric Floors Could Generate Renewable Energy From Walking	Electric Floors Could Generate Renewable Energy From Walking Posted on Oct 29, 2016 Chuck Coker / CC BY-ND 2.0	TRUE
Truthdigger of the Week: Standing Rock Sioux Chairman David Archambault II	Truthdigger of the Week: Standing Rock Sioux Chairman David Archambault II Posted on Oct 29, 2016 Dave Archambault,	TRUE
What the WikiLeaks Revelations Reveal About Donna Brazile and the DNC	What the WikiLeaks Revelations Reveal About Donna Brazile and the DNC Posted on Oct 28, 2016 ( Tim Pierce / CC 2.0 )	FALSE
Tom Hayden Fought for the Rights of All Living Things	Tom Hayden Fought for the Rights of All Living Things Posted on Oct 28, 2016 ( Steve W. Grayson / AP )	FALSE
Here Are Six "Miracle" Drugs Big Pharma Now Regrets - Truthdig	Here Are Six "Miracle" Drugs Big Pharma Now Regrets Posted on Oct 29, 2016	TRUE
Live at Truthdig: Robert Scheer and Thomas Frank Talk About His New Book at 4 P.M.	Live at Truthdig: Robert Scheer and Thomas Frank Talk About His New Book at 4 P.M. PDT Posted on Oct 30, 2016	FALSE
A Trump Named Sue (Video)	A Trump Named Sue (Video) Posted on Oct 30, 2016	FALSE
Now More Than Ever, We Must Tell the Truth About the Iraq War	We Must Tell the Truth About the Iraq War Posted on Oct 30, 2016	TRUE
10 Things Anthony Weiner Did That Were Worse Than Storing Hillary Clinton's Emails on His Computer	10 Things Anthony Weiner Did That Were Worse Than Storing Hillary Clinton's Emails on His Computer Posted on	TRUE
Masoud Barzani Pledges That Kurdish Forces Won't Enter Mosul	Masoud Barzani Pledges That Kurdish Forces Won't Enter Mosul Posted on Oct 29, 2016	FALSE

**Fig. 3** Snapshot of Kaggle’s dataset (KD)

$$\text{Recall} = \frac{TP}{TP + FN} * (100) \tag{4}$$

$$F1 \text{ Score} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} * (100) \tag{5}$$

where TP is true positive, TN is true negative, FN is false positive, and FN is false negative.

### 5.1 Dataset details

Table 1 summarizes the details of two publicly available datasets (Kaggle 2017; Fakenewschallenge 2017) and two self-compiled datasets.

#### A) Kaggle dataset (KD)

Kaggle’s dataset (Kaggle 2017) is a publicly available dataset, shared by Megan Risdal. This text-based dataset consists of news articles with their content and heading. Out

of total 14,565 entries, 6850 are of fake news, and 7715 are of real news. Figure 3 shows a snapshot of Kaggle’s dataset.

#### B) Fake news challenge dataset (FNCD)

Fake news challenge dataset (Fakenewschallenge 2017) is a publicly available dataset. It consists of 28,861 entries out of which 17,168 are of fake news, and 11,693 are of real news. Figure 4 shows a snapshot of fake news challenge dataset.

#### C) Text dataset (TD)

To check the robustness of the proposed model, a self-compiled text dataset of news articles, scraped from the web has been created. This dataset also contains a heading and a content column for each article. It consists of 12,689 entries out of which 5689 are of fake, and 7000 are real. Figure 5 shows a snapshot of this dataset.

Heading	Label
Ferguson riots: pregnant woman loses eye after cops fire BEAN BAG round through car window	TRUE
Crazy Conservatives are sure a Gitmo Detainee killed James Foley	TRUE
A Russian Guy Says His Justin Bieber Rington Saved Him From A Bear Attack	TRUE
Zombie Cat: Buried Kitty Believed Dead, Meows Back to Life	FALSE
Argentina's President Adopts Boy to end Werewolf Curse	TRUE
Next-generation Apple iPhones features leaked	TRUE
Saudi national airlines may introduce gender segregation on its flights	FALSE
Zombie Cat' Claws Way Out Of Grave And into our hearts	FALSE
ISIS might be harvesting organs ,Iraq tells UN	FALSE
Woman has surgery to get third breast: The three most statistically unlikely pairs of boobs	FALSE
EXCLUSIVE: Apple to Unveil The Long-Awaited Retina MacBook Air At its ace Spring Forwarda Event	TRUE
50 foot crab: is this a Crabzilla spotted in the UK?	FALSE
Found! Missing Afghan Soldiers Spotted Trying to Enter Canada	TRUE
Obama orders Fed To Adopt Euro Currency	FALSE
Report: Three missing Afghan Soldiers caught at canadian border	TRUE
<b>(a) Snapshot of headline dataset</b>	
Article Body	Label
Hewlett Packard plans to split into two companies on Monday,according to a report in The wall street	TRUE
Hong Kong protesters have "emulated" the Ferguson gestures in their recent protests	TRUE
Apple has multiple versions of its upcoming smartwatch. This was already know. And while we	TRUE
"My death needs to mean something. My death needs to be counted in the number of transgender people	FALSE
A top iraqi diplomat told world leaders that ISIS is harvesting the organs of its victims to found it murderous	FALSE
The woman claiming to have three boobs is a FRAUD ...A Fake.....and icroically what exposed her lie is a crime	TRUE
ISIS fighters are reportedly just one mile away from Baghdad as reports emerge of al-Qaeda militants	TRUE
In a day packed with major Apple announcements, the company tackled smartphones ,wearables and	TRUE
Islamic state terrorists couldn't strike ransom deals with the United States to trade hostages for money	FALSE
A Five Year old boy has been billed with a "no show fee" for missing his friends BIRTHDAY party	FALSE
This is the stuff nightmares are made of ,but one teen took it all her stride. Walking up in the middle of	FALSE
Saudi Arabia's national airline plans to keep men and women apart after passangers objected to	FALSE
<b>(b) Snapshot of Content Dataset</b>	

Fig. 4 Fake news challenge dataset(FNCD)

A	B	C
Heading	Content	Label
Charlie Gard: Pope and Trump offer parents support	The Vatican said the Pope was following the case "with affection and sadne	TRUE
Jaden Smith STILL not dead after vile 'suicide' hoax continues to baffle	Jaden Smith is alive and well, but his fans are still being fooled by a vile hoax	FALSE
WhatsApp about to start charging you TODAY? Message sparks daily r	Yesterday we brought you the news that Facebook users were being sent a	FALSE
Black Lives Matter supporter didn't kill 19 white women	A racially charged fake news story about a black man killing white women a	FALSE
Modi in Israel: All that happened during PM's historic trip	Narendra Modi's Israel tour was historic because he is the first Indian P	TRUE
FALSE: Obama Signs Executive Order Banning The Pledge Of Allegiance	President Obama has issued an executive order banning the Pledge of Allegi	FALSE
Buckingham Palace denies reports Prince Philip, Queen Elizabeth have	Buckingham Palace has denied reports that the Queen or Prince Philip have	TRUE
FALSE: Justin Trudeau Bans Donald Trump from Entering Canada	Donald Trump has been banned from entering Canada. On 29 December 20	FALSE
Latest Arcade Fire "fake news" story jokes about "removable jihadi be	Arcade Fire have released their latest "fake news" story. These stories	FALSE
Russian hackers 'planted fake news in Qatar to fuel current Gulf crisis,	The FBI and Britain's National Crime Agency (NCA) have deployed officials to	TRUE
Denzel Washington endorsed Trump, according to AmericaNews, Breit	When Ben Collins says "this post of it alone," he is referring to the fact tha	TRUE
FAKE: Atlanta officer did not kill Latoya Smith	ATLANTA - One of the most searched stories on Google has a lot of people t	FALSE
President Trump Launches Twitter Attack: 'Fake News Is the Enemy'	Just hours after President Trump touched down in the United States from a	TRUE

Fig. 5 Snapshot of text dataset (TD)



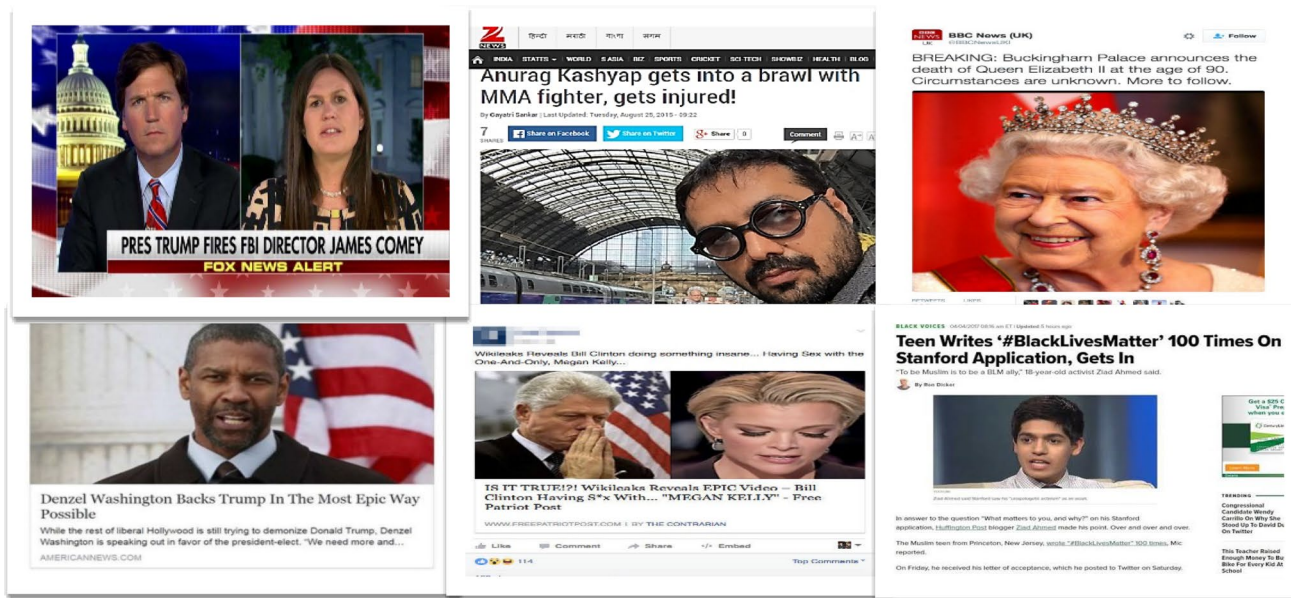


Fig. 6 Snapshot of multimedia dataset (MD)

Table 2 Headings model accuracy (HMA) (%) on three datasets (KD, FNCD, TD)

Dataset	KD	FNCD	TD
HMA (%)	94.16	93.65	91.32

Table 3 Comparison of content models accuracy (%) on the three datasets (KD, FNCD, TD)

Datasets/Parameters	KD	FNCD	TD
Fake news accuracy (FNA)	77.18	79.97	82.45
Real news accuracy (RNA)	85.98	86.15	87.64
Overall accuracy (OAA)	81.85	82.47	85.32

D) Multimedia dataset (MD)

In order to compare the accuracy of the proposed system with state-of-the-art results, multimedia (images) dataset is compiled from microblogs sites (Facebook and Twitter). This dataset consists of 15,664 images with 8496 real news event images and 7168 fake news event image. Figure 6 shows a sample of this dataset.

5.2 Evaluation results and state-of-the-art comparison

The headings model and content models are evaluated separately on the three text-based datasets (KD, FNCD, and TD). Table 2 gives the accuracy achieved on these datasets

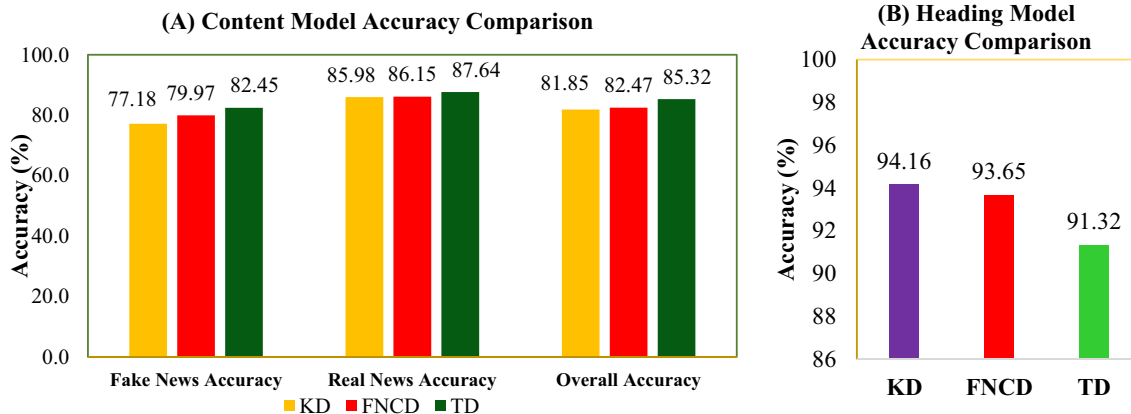


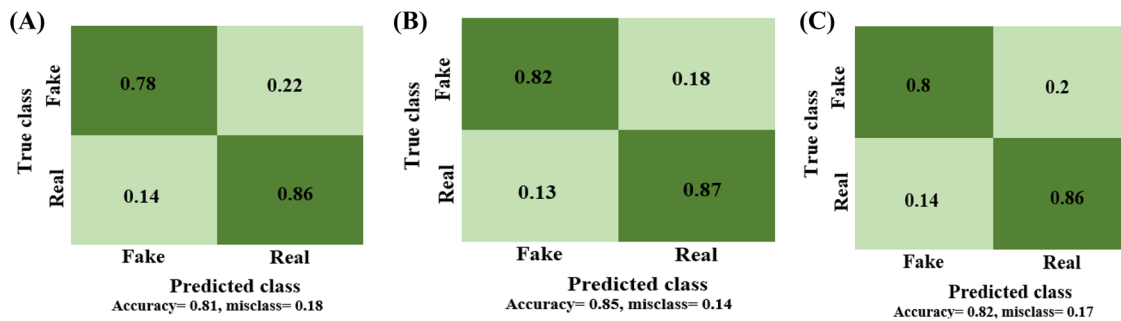
Fig. 7 A Content model accuracy comparison and B Heading model accuracy comparison

**Table 4** Confusion matrix of the content model on three datasets (KD, FNCD, TD)

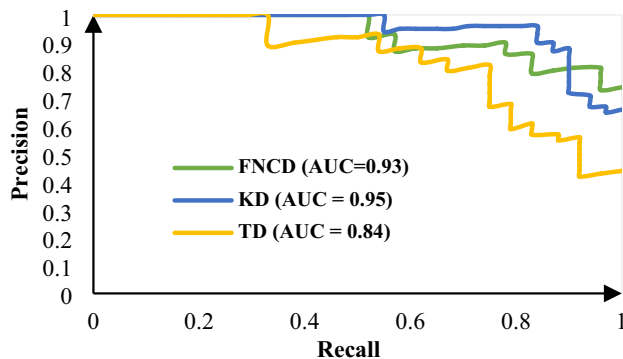
Datasets/Predictions	KD	FNCD	TD
TP	6634	10,073	6135
TN	5287	13,730	4691
FP	1563	3438	998
FN	1081	1620	865

**Table 5** Performance comparison (%) of content models on three datasets (KD, FNCD, TD)

Datasets/Parameters	KD	FNCD	TD
Accuracy	81.85	82.47	85.32
Precision	80.93	74.55	86.00
Recall	85.99	86.14	87.64
F1-score	83.38	79.92	86.81



**Fig. 8** Confusion matrix of the content model on three datasets **A** KD dataset **B** FNCD dataset **C** TD dataset



**Fig. 9** Precision recall curves of the datasets

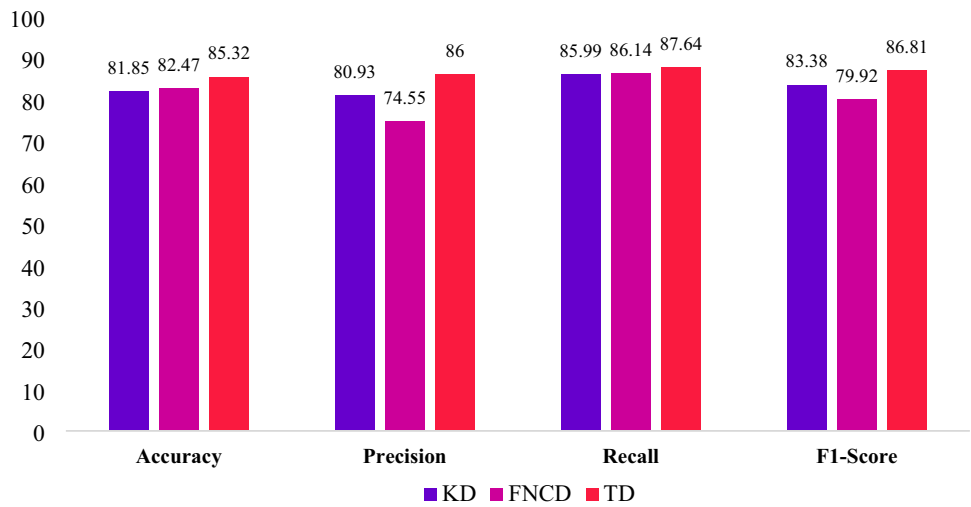
for Headings part of each dataset when passed through our model. The best accuracy of 94.16% is obtained on KD, followed by 93.65% on FD and 91.32% on TD dataset. The detection accuracy is calculated by Eq. 2 as discussed above.

The accuracy of the content model is also computed on three datasets (KD, FNCD, TD). The content of each dataset, when passed through our model, gave the best accuracy of 85.32% on TD dataset. The content model was further tested for three accuracies namely: (1) fake news accuracy, which is the total percentage of fake news that is correctly classified by the model, (2) real news accuracy, which is the total percentage of real news that is correctly classified by

the model, and (3) overall accuracy, which is accuracy calculated using Eq. 2. The best fake news accuracy achieved on the TD dataset is 82.45%. Both KD and FNCD dataset achieved an accuracy of 77.18% and 79.97%, respectively. Best real news accuracy is obtained on the TD dataset with 87.64%, followed by 86.15% on FNCD and 85.98% on KD dataset. Further, the TD dataset achieves an overall accuracy of 85.32% as compared to all the datasets. Table 3 shows the accuracy of the content model on each of the text-based datasets and Fig. 7 shows the comparison chart of various accuracies achieved by the content and heading model on the three datasets. The content model is evaluated using fake news accuracy (FNA), real news accuracy (RNA) and overall accuracy (OAA). Table 4 and Fig. 8 show the confusion matrix of the content model on the three text-based datasets. Since the model has given an acceptable amount of accuracy, it is safe to say that there are structural/linguistic similarities in the content of the articles.

Figure 9. shows the precision-recall curves of the datasets which helps to capture the noise that results from the imbalanced classes. The area under the curve (AUC) helps to compare these curves (Yadav and Vishwakarma 2021). The other evaluation metrics calculated are precision, recall, and F1-score, which are summarized in Table 5. The highest measured values of accuracy, precision, recall, and F1-score are 85.32%, 86.00%, 87.64%, and 86.81%, respectively, for TD dataset. The comparative chart of the content model on the three datasets (KD, FNCD, TD) based on all the evaluation metrics is shown in Fig. 10. Hence, we can say that the

**Fig. 10** Performance comparison chart of the content model on three datasets



**Table 6** Accuracy (%) comparison with similar state-of-the-art models on KD

Method	Experimental settings	Classifier	Accuracy (%)
Yang et al. (1806)	80% training,10% validation,10% testing with fivefold cross validation	CNN	92
O'Brien et al. (2018)	80% training,10% validation,10% testing with fivefold cross validation	CNN	93.5
Samir Bajaj (2017c)	60% training,20% validation,20% testing with fivefold cross validation	LSTM	93
Meel and Vishwakarma (2021)	80% training,10% validation,10% testing with fivefold cross validation	Bi LSTM	92.5
<b>Our model</b>	80% training,10% validation,10% testing with fivefold cross validation	CNN	<b>94</b>

**Table 7** Accuracy (%) comparison with similar state-of-the-art models on FNCD

Method	Classifier	Accuracy (%)
Bhatt et al. (1712)	TF-IDF, BoW, Multi-layer perceptron	89.29
Zhang et al. (2018)	TF-IDF, Deep multi-layer perceptron	86.66
Thorne et al. (2017)	MLP, Decision tree, logistic regression combined in stacked ensemble technique	90.89
Meel and Vishwakarma (2021)	Bi-LSTM	91.65
Meel et. al (2020)	Ensemble of LR, SVM and KNN	92.5
<b>Our model</b>	CNN	<b>93.65</b>

**Table 8** Accuracy (%) comparison with similar state-of-the-art models on TD and MD

Method	Input data format	Classifier	Accuracy (%)
Pathak et al. (2019)	Text	BiLSTM+ Char embedding	90.3
Esmailzadeh et al. (1904)	Text	LSTM	88.78
Ahmed et al. (2018)	Text	LSVM	90
Ajao et al. (2018)	Text	LSTM	81.67
Meel and Vishwakarma (2021)	Text	Bi-LSTM	84.6
Vishwakarma et al. (2019)	Image to text	Rule-based	82
<b>Our method (TD)</b>	<b>Text</b>	<b>CNN</b>	<b>91.32</b>
<b>Our method (MD)</b>	<b>Image to text</b>	<b>CNN</b>	<b>85.06</b>

developed model gives encouraging results for fake news detection.

Further, the proposed model is also implemented on multimedia dataset (MD), and it achieved an accuracy of 85.06%. In order to evaluate the effectiveness of the proposed algorithm, the highest accuracy obtained on the four datasets is compared with the state-of-the-art results.

Tables 6, 7, and 8 compare the proposed model with the state-of-the-art results on KD, FNCD, TD, and MD datasets, respectively. The results obtained by our model are marked in *Bold*. Table 6 demonstrate the comparison of our works with other similar works on KD dataset where experimental settings are given as follows: 80% training set, 10% testing set, and 10% validation set for CNN-based classifier and 60% training set, 20% testing set, and 20% validation set for LSTM-based classifier. Our model gives an accuracy of 94% which is significantly higher than the accuracy obtained by Yang et al. (1806); O'Brien et al. 2018; Bajaj 2017c). This is because in the previous work (Yang et al. 1806), the authors were not able to capture the relevance between headline text and the content text of the articles. Similarly, (Bajaj 2017c) did not consider the heading text of the article for training the model.

Similarly, Table 7 demonstrates the comparison of our work with other similar works on FNCD. (Bhatt et al. 1712) uses feature extraction techniques like term-frequency inverse document frequency (TF-IDF) and bag of words (BoW) followed by multi-layer perceptron. (Zhang et al. 2018) applies deep multi-layer perceptron with TF-IDF and obtained an accuracy of 86.66%. (Thorne et al. 2017) used stacked ensemble classifiers for fake news detection. The main disadvantage of Bhatt et al. (1712); Zhang et al. 2018; Thorne et al. 2017) is that they just classified the stance of user comments or tweets on some news post as agree, disagree, discuss and unrelated which is just a subpart of fake news detection. The algorithms in all the above stated works did not provide the overall credibility status of a news as genuine or fake. Hence, from Table 6 and Table 7, it can be said that our model performs better than the existing approaches. Thus, we see a significant improvement in accuracy over the earlier works. This extra accuracy has much more significance, as the effect of fake news that is being consumed by a user is severe.

Most of the previous work has focused on textual data. Since the TD and MD datasets are self-compiled datasets, hence to test the effectiveness of the proposed model we have compared the state-of-the-art techniques by evaluating them on the TD and MD datasets. The comparison results are shown in Table 8. Our model shows the highest accuracy of 91.32% on TD dataset and 85.06% on MD dataset. (Fakenewschallenge 2017) applied BiLSTM with character embedding but did not perform any pre-processing on the data. (Esmailzadeh et al. 1904; Ajao et al.

2018) used LSTM for classifying the news into fake or real. However, the hybrid approach of LSTM-CNN in Ajao et al. (2018) increased the accuracy by around 10%, which demonstrates the power of CNN for classification; still this hybrid approach proved to be computationally expensive. (Yang et al. 1806) applied linear SVM (LSVM) for classification but didn't consider the features that reflect the writer's style.

The above experimental results show that the proposed model for fake news detection on text and multimedia dataset is quite satisfactory. The accuracy over all the four datasets Kaggle's dataset, fake news challenge dataset, self-compiled text dataset, and the multimedia dataset is a good improvement over the state-of-the-art models. Web scraping and removal of low credibility links using verified source model (VSM) are the two salient features that have helped in improving the accuracy of our results. Finally, CNN model applied on heading and body part of news text exploits the inbuilt linguistic patterns of fake news that helped in segregating it from real news.

## 6 Conclusion

Fake news has a powerful, pervasive, and persistent force which is in general circulation without confirmation or certainty of facts. Fake news spread quickly contributing to widespread chaos and sensation in the absence of first-hand information from traditional sources during an emergency crisis. In the present work, we have tried to solve the problem of news verification with the help of WSCH-CNN model, built on top of word2vec, that proved quite successful in detecting the fake news with a high accuracy of 85.06% on multimedia dataset, 94.16% for heading model and 85.32% for content model which supersedes the state-of-the-art. The experimental results suggest that there exists a linguistic pattern between fake news as fake news writers opt a particular style of writing to persuade readers to believe them as real. These linguistic patterns can be successfully analyzed to recognize fake news, as there are a limited number of fake content producers.

This work provides a novel framework of fake news detection using four major components processing unit, verified source model, CNN model and decision-making model. The efficacy of the proposed framework is being verified on four different datasets and state-of-the-art baseline comparisons. Further, the work can be extended for multimedia audio-visual data as social media is becoming increasingly vulnerable toward fake news spread in the form of videos. Unsupervised and semi-supervised fake news detection technologies are a potential future line of research as with every second of time social media is producing huge volumes of unannotated real-time data to be checked for credibility.

Cross-platform spreading of misinformation can be another major field of interest for researchers.

**Author contributions** Dinesh Kumar Vishwakarma and Priyanka Meel conducted the experiment and written the manuscript. Ashima Yadav and Kuldeep Singh done the proof read of the manuscript.

## Declarations

**Conflict of interest** The authors declare no competing interests.

**Conflict of interest** All authors declare that they have no conflict of interest.

## References

- Ahmed H, Traore I, Saad S (2018) "Detecting opinion spams and fake news using text classification. *Secur Priv* 1(1):e9
- Aiello LM, Petkos G, Martin C, Corney D, Papadopoulos S, Skraba R, Goker A, Kompatsiaris Y, Jaimes A (2013) Sensing trending topics in twitter. *IEEE Trans Multimed* 15(6):1268–1282
- Ajao O, Bhowmik D and Zargari S (2018) Fake news identification on twitter with hybrid CNN and RNN models. In: 9th International Conference on Social Media and Society, Copenhagen
- Alkhodair SA, Ding SH, Fung BC and Liu J (2020) Detecting breaking news rumors of emerging topics in social media. *Inf Process Manag* 57(2):102018
- Alrubaian M, Al-Qurishi M, Hassan MM, Alamri A (2018) A credibility analysis system for assessing information on twitter. *IEEE Trans Depend Secure Comput* 15(4):661–674
- Bajaj S (2017c) The Pope Has a New Baby! Stanford University
- Barbado R, Araque O, Iglesias CA (2019) A framework for fake review detection in online consumer electronics retailers. *Inf Process Manag* 56(4):1234–1244
- Bhatt G, Sharma A, Sharma S, Nagpal A, Raman B and Mittal A (2017) On the benefit of combining neural, statistical and external features for fake news identification. arXiv preprint: [arXiv:1712.03935](https://arxiv.org/abs/1712.03935)
- Bondielli A, Marcelloni F (2019) A survey on fake news and rumour detection techniques. *Inf Sci* 497:38–55
- Buntain C and Golbeck J (2017a) Automatically identifying fake news in popular twitter threads. In: IEEE International Conference on Smart Cloud, New York
- Campan A, Cuzzocrea A and Truta TM (2017) Fighting fake news spread in online social networks: actual trends and future research directions. In: IEEE International Conference on Big Data (Big Data), Boston
- Cheung M, She J, Jie Z (2015) Connection discovery using big data of user-shared images in social media. *IEEE Trans Multimed* 17:1417–1428
- Chidiac NM, Damien P and Yaacoub C (2016) A robust algorithm for text extraction from images. In: 39th International Conference on Telecommunications and Signal Processing, Vienna
- Conneau A, Schwenk H, Barrault L and Lecun Y (2016) Very deep convolutional networks for text classification. In: Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics, Valencia, Spain
- Egele M, Stringhini G, Kruegel C, Vigna G (2017) Towards detecting compromised accounts on social networks. *IEEE Trans Depend Secure Comput* 14(4):447–460
- Elkasrawi S, Dengel A, Abdelsamad A and Bukhari SS (2016) What you see is what you get? Automatic image verification for online news content. In: 12th IAPR Workshop on Document Analysis Systems (DAS), Santorini, Greece
- Esmaeilzadeh S, Peh GX and Xu A (2019) Neural abstractive text summarization and fake news detection. arXiv preprint: [arXiv:1904.00788](https://arxiv.org/abs/1904.00788)
- Fakenewschallenge (2017) "News dataset," [Online]. Available: <https://www.fakenewschallenge.org>.
- Figueira Á, Oliveira L (2017) The current state of fake news: challenges and opportunities. *Proc Comput Sci* 121:817–825
- Garg G and Sharma D (2020) Phony news detection using machine learning and deep-learning techniques. In: 9th International Conference System Modeling and Advancement in Research Trends (SMART), Moradabad, India, pp 27–32
- Gilda S (2017b) Evaluating machine learning algorithms for fake news detection. In: IEEE 15th Student Conference on Research and Development, Putrajaya
- Granik M and Mesyura V (2017) Fake news detection using naive Bayes classifier. In: IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON), Kiev
- Jang SM, Geng T, Li JQ, Xia R, Huang C, Kim H, Tang J (2018) A computational approach for examining the roots and spreading patterns of fake news: evolution tree analysis. *Comput Hum Behav* 84:103–113
- Jhonson R and Zhang T (2015) Effective use of word order for text categorization with convolutional neural networks. In: The 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Colorado
- Jin Z, Cao J, Zhang Y, Zhou J, Tian Q (2017) Novel visual and statistical image features for microblogs news verification. *IEEE Trans Multimed* 19(3):598–608
- Kaggle (2017) "Kaggle fake news dataset," [Online]. Available: <https://www.kaggle.com/mrisdal/fake-news>.
- Kim Y (2014) Convolutional neural networks for sentence classification. In: Proceedings of the 2014 conference on Empirical Methods in Natural Language Processing, Doha, Qatar
- Liu Y, Jin X, Shen H (2019) Towards early identification of online rumors based on long short term. *Inf Process Manag* 56(4):1457–1467
- Meel P, Chawla P, Jain S and Rai U (2020) Web text content credibility analysis using max voting and stacking ensemble classifiers. In: Advanced Computing and Communication Technologies for High Performance Applications (ACCTHPA), Kerala
- Meel P and Vishwakarma DK (2021) Deep neural architecture for veracity analysis of multimodal online information. In: 11th International Conference on Cloud Computing, Data Science & Engineering, Noida
- Meel P, Vishwakarma DK (2019) Fake news, rumor, information pollution in social media and web: a contemporary survey of state-of-the-arts, challenges and opportunities. *Expert Syst Appl* 153:112986
- Nisar TM, Prabhakar G, Strakova L (2019) Social media information benefits, knowledge management and smart organizations. *J Bus Res* 94:264–272
- Niu X (2008) On two major issues in the current network news communication. In: IEEE International Symposium on IT in Medicine and Education, Xiamen, China
- O'Brien N, Latessa S, Evangelopoulos G and Boix X (2018) The language of fake news: opening the black-box of deep learning based detectors In: 32nd Conference on Neural Information Processing Systems (NIPS 2018), Montréal
- Pathak A and Srihari RK (2019) BREAKING! presenting fake news corpus for automated fact checking. In: Proceedings of the 57th

- Annual Meeting of the Association for Computational Linguistics: Student Research Workshop
- Shaabani E, Guo R and Shakarian P (2018) Detecting pathogenic social media accounts without content or network structure. In: International Conference on Data Intelligence and Security, South Padre Island
- Sharma K, Qian F, Jiang H, Ruchansky N (2019) Combating fake news: a survey on identification and mitigation techniques. *ACM Trans Intell Syst Technol* 10(3):1
- Sharma D, Singh B, Agarwal S, Kim H, Sharma R (2022) Sarcasm detection over social media platforms using hybrid auto-encoder-based model. *Electronics* 11:2844
- Shen Y, He X, Gao J, Deng L and Mesnil G (2014) Learning semantic representations using convolutional neural networks for web search. In: Proceedings of the 23rd International Conference on World Wide Web, Seoul
- Shu K, Wang S and Liu H (2018) Understanding user profiles on social media for fake news detection. In: IEEE Conference on Multimedia Information Processing and Retrieval, Miami
- Singh B and Sharma D (2021b) Image forgery over social media platforms—a deep learning approach for its detection and localization. In: INDIACOM-2021b - International Conference on computing for Sustainable Global development – Bharti Vidyaapeeth, Delhi
- Singh B and Sharma D (2021) Detecting image forgery over social media using residual neural network. In: International Conference on Artificial Intelligence and Sustainable Engineering (AISE-2020)
- Singh B, Sharma D (2021a) SiteForge: detecting and localizing forged images on microblogging platforms using deep convolutional neural network. *Comput Ind Eng* 162:107733
- Singh B, Sharma D (2022) Predicting image credibility in fake news over social media using multi-modal approach. *Neural Comput Appl* 34:21503–21517
- Tang S, Blenn N, Doerr C, Miegheem PV (2011) Digging in the digg social news website. *IEEE Trans Multimed* 13:1163–1175
- Thorne J, Chen M, Myrianthous G, Pu J, Wang X and Vlachos A (2017) Fake news stance detection using stacked ensemble of classifiers. In: Proceedings of the 2017 EMNLP Workshop: Natural Language Processing meets Journalism
- Vicario MD, Quattrocioni W, Scala A, Zollo F (2019) Polarization and fake news: early warning of potential misinformation targets. *ACM Trans Web* 13(2):1–22
- Vishwakarma DK, Varshney D, Yadav A (2019) Detection and veracity analysis of fake news via scrapping and authenticating the web search. *Cogn Syst Res* 58:217–229
- Viviani M, Pasi G (2017) Credibility in social media: opinions, news, and health information—a survey. *Wiley Interdiscip Rev Data Min Knowl Discov* 7(5):6–25
- Wessel M, Thies F, Benlian A (2016) The emergence and effects of fake social information: evidence. *Decis Support Syst* 90:75–85
- Yadav A, Vishwakarma DK (2021) A deep multi-level attentive network for multimodal sentiment analysis. *ACM Trans Multimed Comput Commun Appl* 28:1–30
- Yang Y, Zheng L, Zhang J, Cui Q, Zhang X, Li Z and Yu PS (2018) TI-CNN: convolutional neural networks for fake news detection. arXiv preprint: [arXiv:1806.00749](https://arxiv.org/abs/1806.00749)
- Yih W, He X and Meek C (2014) Semantic parsing for single-relation question answering. In: Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics, Maryland
- Zhang Q, Yilmaz E and Liang S (2018) Ranking-based method for news stance detection. In: International World Wide Web Conferences Steering Committee
- Zubiaga A, Aker A, Bontcheva K, Liakata M, Procter R (2018) Detection and resolution of rumours in social media: a survey. *ACM Comput Surv* 51(2):1–36

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.