Check for updates

Received: 25 April 2019 | Revised: 31 July 2019 | Accepted: 2 September 2019

DOI: 10.1002/ett.3767

RESEARCH ARTICLE

WILEY

Fake news detection using deep learning models: A novel approach

Sachin Kumar[1] | **Rohan Asthana[1]** | **Shashwat Upadhyay[2]** | **Nidhi Upreti[1]** | **Mohammad Akbar[1]**

[1]Department of Computer Science and Engineering, Ajay Kumar Garg Engineering College, Ghaziabad, India

[2]Department of Information Technology, Ajay Kumar Garg Engineering College, Ghaziabad, India

Correspondence
Sachin Kumar, Department of Computer Science and Engineering, Ajay Kumar Garg Engineering College, Ghaziabad-201 009, India.
Email: imsachingupta@rediffmail.com

Abstract

With the ever increase in social media usage, it has become necessary to combat the spread of false information and decrease the reliance of information retrieval from such sources. Social platforms are under constant pressure to come up with efficient methods to solve this problem because users' interaction with fake and unreliable news leads to its spread at an individual level. This spreading of misinformation adversely affects the perception about an important activity, and as such, it needs to be dealt with using a modern approach. In this paper, we collect 1356 news instances from various users via Twitter and media sources such as PolitiFact and create several datasets for the real and the fake news stories. Our study compares multiple state-of-the-art approaches such as convolutional neural networks (CNNs), long short-term memories (LSTMs), ensemble methods, and attention mechanisms. We conclude that CNN + bidirectional LSTM ensembled network with attention mechanism achieved the highest accuracy of 88.78%, whereas Ko et al tackled the fake news identification problem and achieved a detection rate of 85%.

1 | INTRODUCTION

Online information is always at one's disposal because of being just a few clicks away. With the creative freedom given to users to share a story, the difficulty to determine the root of false information grows day by day. The presence of clickbait titles and dramatic headlines is at its peak, which aids in the propagation of unprofessional and inaccurate news in return for ad revenue. Users, wanting to be part of such a hot topic/discussion, modify the original message by mistake or with intention, which ultimately leads to the diffusion of rumor in the network.[1]

Fake news is written with the intent to spread information disguised as propaganda or a hoax that results in financial or political gain,[2] one which may be used to sway public opinion toward falsity. This even persuades people's ideologies and their beliefs to a larger extent, which may cause a lot of harm. Such persuasion is prominent when a major news story breaks out, where the supporters generally tend to share information in its full originality, whereas the ones whose opinions do not align with the said information resort to sharing that same information with some modifications of their own.

In the present scenario, media outlets are not the only source of information. Individual involvement in news sharing has grown significantly over the past few years to a level where it is becoming increasingly difficult to differentiate news that originates from a credible source from the one that is fabricated. As a result, fake news has received a lot of research attention in the recent years by organizations such as Facebook, Google, Twitter, and by many researchers, who are putting constant efforts into combating the spread of fabricated stories.

 © 2019 John Wiley & Sons, Ltd.

In a study, researchers at Stanford found that students have difficulty in determining the credibility of information online.[3] Furthermore, poor journalism leads to misinterpretation of the actual news itself due to false connection, misleading content, and false context. This improper communication on their part generally makes misinformation creep into the real news system, usually with loss of original and factual information.[4] Other instances are those articles which, when read by the reader, may undergo modifications as per the user's sentiment toward that topic. Due to these subsequent modifications, the news deviates from its initial meaning.

2 | LITERATURE REVIEW

The biggest contributing factor to the distribution of fake news is the social media as an independent platform. Propagation of fake news can be carried out via bots or by users themselves.

Bots are algorithms that execute specific functions using the corresponding input and the related response patterns. Since the users mostly believe whatever they read on the internet, so bots spam fake news in high volumes to make it look convincing.[5] In addition, users are susceptible to sharing those stories, which tend to have more likes or comments on them because it generally receives more attention than other stories. Emotions and feelings toward a topic of interest play yet another role in deciding a user's likeness of a topic.[5] As a result, false news diffuses faster than original news because humans are involved in this activity as much as the bots, according to the researchers at MIT.[6]

This research also showcased the propagation of false claims and stories on a larger scale and peaked during major events (Figure 1).

2.1 | Tackling fake news

Fake news is not a new phenomenon. In fact, this issue has been tackled ever since its rise has caused adverse effects both in the technical as well as the political fields. It is necessary to address it because the day-to-day dependence on social networking sites to get news stories from is on a constant rise due to technological availability and it is showing no signs of declining any time soon. The tech giant Facebook has already taken steps to combat the circulation of misinformation on its website (in some countries) by partnering with third-party fact-checkers to review articles and posts on Facebook and rate their accuracy.[7] By identifying the false news content, it gets pushed down in the news feed and action is taken against repeat offenders.

Previous tactics to stop the diffusion of false information have largely focussed their research toward the fake news stories that are spread by bots.[8] Bots, generally, are accounts having a social media presence that tend to share misinformation more often than the genuine accounts, ie, the frequency with which the bots spread irrelevant news is seemingly high. Their prime targets are those users who tend to show some influence on others (particularly their social media followers). It was also noticed that the bots that were used to spread false information to the target audience were usually active in the early phases of the fake news propagation and as such, they attracted humans of similar thinking who reposted the same news in their feed.[8] It has similarly been found that social bots often populate the social space to cause

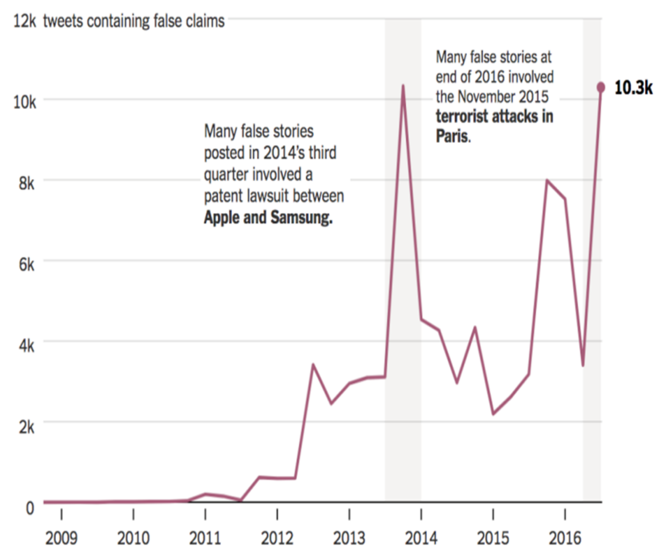

FIGURE 1 Analysis by The New York Times. Source: MIT Media Lab[6]

harm and deceive social media users. They have also been used to cause political disruptions, negatively affect the stock market, for theft of personal information, and to spread misinformation.[9]

The main issue arises with the dynamic nature of the network. Since the network deals with real-time data, it is necessary to control the diffusion of rumor early in the process.[10] One of the approaches to stop the dissemination of rumor was the identification of its source.[11] The identification of a proper diffusion model led to the analysis about the pace at which the dissemination of fake news occurs. Then, based on the source, either the anti-rumor–based approach (for a single source) or the approximation-based approach (for multiple sources) was employed to tackle the spread of rumor in the network.

2.2 | Previous approach

Several algorithms have been tested to best detect the misleading stories from nonreputable sources. These algorithms, that have attained accuracies better than previous methods, include the following:

1. The detection of fake news using natural language processing (NLP) techniques by applying term frequency-inverse document frequency (TF-IDF) of bi-grams and probabilistic context-free grammar detection to news articles.[12] The dataset was tested on multiple classification algorithms such as support vector machines (SVMs), stochastic gradient descent, gradient boosting, bounded decision trees, and random forests and it was found that TF-IDF of bi-grams fed into a Stochastic Gradient Descent model identifies noncredible sources with an accuracy of 77.2%.
2. A CSI model that combines the three generally agreed upon characteristics of fake news, ie, the text of an article, the user response it receives, and the source users promoting it, for a more accurate and automated prediction.[13] This model is composed of the following three modules: capture, score, and integrate. The first module is based on the response and text that employs the use of a recurrent neural network (RNN) to capture the temporal pattern of user activity. The second one learns the source characteristic based on the users' behavior, and then the third module integrates both of them to classify whether an article is fake or not.
3. Fake news detection using naive Bayes classifier[14] that achieved classification accuracy of approximately 74%, and whose results suggest that fake news detection problem can be addressed with artificial intelligence methods.

2.3 | Comparative study with backtracking methodology based on cognitive system

Ko et al[15] defined the possibility of news being fake using the reverse-tracking methodology of articles posted on the cognitive system. The lack of fake news distinction time compared to the speed of information sharing on the internet makes it difficult to distinguish fake news from the real news. Moreover, when factors such as the diversity of articles and the strong subjectivity of news stories are considered, then fake news detection becomes a difficult task. Using a backtracking approach based on cognitive system, Ko et al tackled the fake news identification problem and achieved a detection rate of 85%.

2.4 | Shortcomings of backtracking methodology based on cognitive system

While the backtracking approach works well in case of known users and online profiles, it falls behind in the aspect when the user's account has either been removed or it has been suspended. In both of these cases, the lack of origin of the information makes detection at the end quite challenging. With the current trend of social media organizations such as Facebook, Twitter, Reddit, etc, striving to constantly improve their users' experience by banning or suspending user accounts due to suspicious activity, it will make it even more difficult to detect fake news using this approach in the long run.

Furthermore, subjective posts by media journalists or freelancers have a hard time being categorized into the fake or real news category because even if the news they share is real, it will still carry with itself the writer's own philosophy added to it, thereby making the detection task challenging.

In this paper, our team proposes models that achieved appreciable accuracy and attained a maximum accuracy of 88.78%, thus considerably improving over the current state-of-the-art methods.

Table 1 and Table 2 are two different datasets that are formed using the same topic. Each dataset consists of thirty tweets and the news instances in each dataset are sorted according to their timestamps for a clear representation.

TABLE 1 Real news dataset 1: Chester Bennington commits suicide

Number	News	Date
1.	Chester Bennington dies of an apparent suicide by hanging at his home on July 20th, 2017. RIP ChesterBennington -We're missing you alot!	20/07/17
2.	Gavin Rossdale reacts onstage to Chester Bennington's suicide.	22/07/17
3.	Chester Bennington cause of death revealed: rocker died of suicide by hanging, coroner confirms.	24/07/17
4.	Chester Bennington's wife Talinda breaks silence one week after singer's suicide.	29/07/17
5.	Linkin Park launched a suicide awareness site in honor of late frontman Chester Bennington.	29/07/17
6.	This is my latest on the passing of Chester Bennington, mental illness, suicide, & the fear of being a burden.	03/08/17
7.	Linkin Park, Ken Jeong taped *Carpool Karaoke* a week before Chester Bennington committed suicide.	12/10/17
8.	Suicide is a preventable death: remembering Chester Bennington from Linkin Park.	12/10/17
9.	This is wrong there are videos of Chester Bennington laughing and having a good time with his family a few days before he committed suicide.	30/10/17
10.	*STONE TEMPLE PILOTS Bassist: CHESTER BENNINGTON's Suicide 'Was A Huge, Huge Surprise.'*	30/10/17
11.	It marked the first time Linkin Park has performed since Chester Bennington died by suicide.	05/11/17
12.	The suicide deaths of rockers Chris Cornell and Chester Bennington inspired a local country artist to put on a free event and share his struggles in a new song.	12/11/17
13.	PAPA ROACH Drummer says news of CHESTER BENNINGTON's suicide *literally* took his breath away.	25/11/17
14.	More details about Chester Bennington's suicide have been revealed.	05/12/17
15.	The story reported by TMZ regarding Chester Bennington's autopsy is abusive and harmful to thousands of Twitter members. Reporting information like this can lead to more deaths by suicide. Please remove immediately!	07/12/17
16.	People genuinely need to start understanding that suicide is not what kills a person who is so sad and not in the right place, mentally. It's depression, which is a silent killer. Chester Bennington was far from *selfish* … I wish people would see the reality of mental illness.	10/12/17
17.	2017 has seen 2 of my favourite main vocalists passed away due to suicide over depression. Chester Bennington and Kim Jonghyun.	18/12/17
18.	When Chester Bennington commited suicide, no one makes fun of it. Then, why can't you do the same to Jonghyun, by giving simple respect & condolences. Regardless of fandom, life is life, no one should deserve this.	18/12/17
19.	This was Chester Bennington a day before his suicide. And this is Jonghyun a week before his death. Similarity? They looked happy even when the demons inside them already ate them up.	19/12/17
20.	Chester Bennington and Avicii deserve so much more than this. I don't understand why TMZ thinks it's okay to post information about someone's death like they do, especially when it's a suicide. This is one reason why mental health and suicide still have a stigma.	01/05/18
21.	It's sad how many artists we've lost to suicide. It's hard to believe Robin Williams, Chris Cornell, Chester Bennington, Avicii, Scott Hutchison, Elliott Smith, David Foster Wallace, Kurt Cobain, Jeff Buckley, and so many others are gone. Call 1-800-273-8255 to talk to someone.	11/05/18
22.	We don't cover the fashion industry, but designer Kate Spade sadly died by suicide today. We in the rock community tragically lost Chris Cornell and Chester Bennington last year. Call 1-800-273-8255 if you ever have suicidal thoughts and need help.	05/06/18
23.	Suicide rates have increased more than 25% since 1999. @mikeshinoda, co-founder of linkinpark, remembers lead singer Chester Bennington in light of the deaths of Kate Spade and Anthony Bourdain	06/09/18
24.	Mike Shinoda reflects on losing bandmate Chester Bennington to suicide and the future of Linkin Park	19/06/18
25.	On Fox 11 News Thursday at 10 PM. Understanding suicide, one year after the death of Linkin Park's Chester Bennington … and his wife's fuel to help others.	18/07/18
26.	Exactly how early is it??? Thank you GDLA for having me. On w/ @Megancolarossi @Elex_Michaelson to talk about my story airing tonight on Fox 11 news at 10 PM FOXLA. Understanding suicide on the one year anniversary of the death of Linkin Park's Chester Bennington.	19/07/18
27.	One year ago, we lost Chester Bennington to suicide. We send our love to his family, friends, and fans today.	20/07/18
28.	Guess who released extensive details of Chester Bennington's death and previous suicide attempt that his wife didn't want his kids to see … TMZ	08/09/18

TABLE 1 (continued)

Number	News	Date
29.	In July 2018, 48 streamers came together to spread light in honor of Chester Bennington and all of those lost to suicide. In six days, five come together again to reiterate our goals. Let's do this.	11/01/19
30.	This goes out to Chester Bennington, Robin Williams, and others that have fallen prey to the Leftist religion of death, abortion, suicide, and hate. The hatred that the left portrays only leads to destruction and death. Let's rise past this, and learn.	13/02/19

Table 1 comprises of real news stories on the topic: *“Chester Bennington commits suicide.”*
Table 2 comprises of fake news stories on the topic: *“Was Chester Bennington murdered?”*
Table 3 and Table 4 on the next page are different datasets that are formed using vastly different topics.
Table 3 comprises of real news stories on the topic: *“Mumbai terror attack.”*
Table 4 comprises of fake news stories on the topic: *“Denzel Washington endorses Donald Trump.”*

3 | OUR METHOD

In this paper, we aim to classify fake news stories from real news stories. Using 1356 news instances, we develop and train several models and obtain their accuracies. Our models for fake news classification are based upon the sentiment analysis of users in the social media. We use the architectures to detect patterns in our data, where patterns can be anything such as unusual capitalization, random exclamations/question marks, etc.

3.1 | Data collection

Our research team worked on collecting datasets of fake and real news using Twitter's Advanced Search functionality. We cover news stories on a variety of topics and prepare real and fake news datasets having 30 tweets in each. In addition to the datasets created by us, we also use the dataset provided by FakeNewsNet,[16-18] which contains 1056 real and fake news samples from PolitiFact (Table 5 and Table 6).

3.2 | Data preprocessing

Due to versatility in formats of different data, it becomes necessary to preprocess and encode the data before feeding it into a network. Our dataset preprocessing involves the following steps: First, we extract each news title from the dataset using BeautifulSoup library in Python[19] and remove “\\”, “’”, and “\’”.

```
for idx in range (politi.title.shape[0]):
    text_politi = BeautifulSoup (politi.title[idx]).
    texts_politi.append (clean_str(str (text_politi.get_text().encode()))).

for idx in politi[‘Class’]:
    label.append (idx).
def clean_str(news):
    news = re.sub(r”\\”, ““, news)
    news = re.sub(r”\’“, ““, news)
    news = re.sub(r”\”“, ““, news)
    return news.strip().
```

Second, we append the encoded version of each news title into a list and use tokenizer to split each title into several tokens.

```
tokenize = Tokenizer (num_words = 20000)
tokenize.fit_on_texts(texts_politi)
sequence = tokenize.texts_to_sequences(texts_politi)
word_indexes = tokenize.word_index
print(‘Number of Unique Tokens’,len (word_indexes)).
```

TABLE 2 Fake news dataset 1: Was Chester Bennington murdered?

Number	News	Date
1.	Police: Chester Bennington was murdered.	21/07/17
2.	Apparently, the Clinton Foundation murdered Chester Bennington.	21/07/17
3.	Chester Bennington is now suspected MURDERED. Was working at a foundation to prevent child sexual abuse| pizzagate.	21/07/17
4.	News just came in apparently Chester Bennington may have been murdered police are looking just in case.	22/07/17
5.	Did Chester Bennington get murdered bc he was about to expose PedoGate?	22/07/17
6.	Shocker: Chester Bennington murdered by dad John Podesta! Police have reportedly launched a murder investigation.	23/07/17
7.	Chester Bennington murdered by John Podesta? Police launch an investigation!!	24/07/17
8.	PIZZAGATE! CHESTER BENNINGTON MURDERED BY DAD JOHN PODESTA!	26/07/17
9.	Chester Bennington murdered by Dad John Podesta!	27/07/17
10.	Chester Bennington's friend: 'He was murdered by Elite Pedophiles.'	28/07/17
11.	*Yes, 100% I believe both Chris Cornell & Chester Bennington were murdered.*	29/08/17
12.	WIKILEAKS DOCUMENT PROVES CHESTER BENNINGTON MURDERED BY JOHN PODESTA PEDOGATE.	31/08/17
13.	Chester Bennington, who were allegedly murdered, as they were close to revealing some dangerous secrets they had uncovered about Hollywood and celebrities involved in these rings.	22/11/17
14.	I'm still questioning the deaths of Chris Cornell and Chester Bennington, who were allegedly murdered.	22/11/17
15.	The internet says Chester Bennington was murdered to cover up a Pizzagate-related conspiracy.	23/11/17
16.	Chester Bennington and Chris Cornell murdered for trying to expose David Geffen and entertainment industry pedophilia ring. Official story both committed suicide.	10/12/17
17.	Chester Bennington was murdered.	02/03/18
18.	Chester Bennington was murdered because he was getting too close to exposing a pedophile ring.	30/03/18
19.	Both Chester Bennington and Chris Cornell were murdered for exposing this … Did you know that they were about to expose a massive Elite #PedoGate ring?	24/04/18
20.	Chester Bennington and Chris Cornell, murdered to prevent pedo secrets?	29/05/18
21.	Chester Bennington was MURDERED!!!!!!!!!!!!	31/05/18
22.	I don't think Anthony Bourdain, Avicii, or Chester Bennington killed themselfs.. If you look into it they were all trying to expose pedophilia in Hollywood so I believe they were murdered to keep them from using their giant platforms to speak out.	18/06/18
23.	Still can't believe digitalmusicnews ran a piece questioning whether or not Chester Bennington was murdered by a violent pedophilia ring.	24/07/18
24.	Chester Bennington was murdered for trying to expose Pizzagate. Really.	24/07/18
25.	Report: CHESTER BENNINGTON – Fans believe he was murdered.	25/07/18
26.	Was Chester Bennington murdered because of this- The Silent Children Documentary	05/09/18
27.	Chester Bennington was going to appear in the documentary silent children which was going to expose child trafficking he was murdered.	27/09/18
28.	You had your son Chester Bennington & his best friend murdered. Yeah we all know.	26/10/18
29.	Let's not forget where he was the night Madeleine McCann disappeared and he had Chester Bennington his biological son murdered.	14/11/18
30.	This is the exact reason Chris Cornell and Chester Bennington were murdered last year. They were both working on a documentary about Hollywood pedophiles.	14/11/18

Third, we pad the sequences using maximum sequence length as 1000. Then, we randomly shuffle our data so that the classifier does not predict the labels according to the news topic itself. The data are split into train and test sets comprising of 60% training data and 40% test data.

data = pad_sequences(sequence, maxlen = 1000)
label = to_categorical(np.asarray (label))
indices_data = np.arange (data.shape[0])

TABLE 3 Real news dataset 2: Mumbai terror attack

Number	News	Date
1.	England cricket team to fly out of India on Friday after militant attack in Mumbai - England cricket board.	27/11/08
2.	WORLD: Terror in India: Mumbai under assault: In the most intense recent attack on an Indian city.	27/11/08
3.	Mumbai gunman's confession in limbo: Court defers decision on accepting Mumbai attack gunman's surprise confession.	21/07/09
4.	Terror ties run deep in Pakistan, Mumbai attack case shows.	27/07/09
5.	Head of group tied to Mumbai attack arrested.	21/09/09
6.	Bring perpetrators of Mumbai attack to justice: Obama, Manmohan.	25/11/09
7.	Parliament remembers Mumbai attack victims.	26/11/09
8.	As Mumbai recalls attack, security concerns persist.	26/11/09
9.	NUS law faculty has set up a scholarship in memory of Ms Lo Hwei Yen who died in the Mumbai terror attack a year ago.	26/11/09
10.	Taj Mahal Palace Mumbai will re-open its magnificently restored Palace Wing on August 12, 2010. It was closed after 26/11 attack.	09/08/10
11.	3 years before deadly '08 Mumbai attack, wife of key figure told feds he had joined terror group, trained in Pakistan.	16/10/10
12.	Mumbai attack: Kasab laughs at hearing.	18/10/10
13.	India ties Pakistani agency to Mumbai attack.	19/10/10
14.	POTUS just visited memorial to the victims of 26/11 terror attack here in Mumbai … Now visiting with families of some of the victims.	16/11/10
15.	Mumbai massacre survivor Will Pike on compensation: *If you are caught up in a terrorist attack abroad there is nothing available to you.*	24/11/10
16.	Mumbai attack victims remembered: Memorials and candlelight march mark second anniversary of deadly attacks.	26/11/10
17.	Indian intelligence had no warning of a possible attack on #Mumbai, Home Minister Chidambaram says #mumbaiblasts.	14/07/11
18.	Electric circuit found on one dead in Mumbai terror attack #mumbaiblast #ht.	14/07/11
19.	Silent candle light march at INDIA GATE, New Delhi, today (14th July 2011) at 7.45 pm in wake of terrorist attack in Mumbai.	14/07/11
20.	No help request from India on Mumbai attack: US.	16/07/11
21.	26/11 Mumbai attack: HR practices converted ordinary Taj employees into heroes.	25/11/11
22.	London High Court to decide if #Mumbai terror attack victim can bring case in the UK.	26/11/13
23.	Police leaflets tell commuters to 'run, hide and tell' in event of Mumbai-style terror attack.	26/11/14
24.	Our colleague Commando Surinder who survived 26/11 attack is attending a peace prayer at Taj Mumbai. Nation must stand united against terror.	26/11/14
25.	Today as we remember Mumbai terror attack, let us work together to combat terrorism: PM Modi addresses SAARC summit.	26/11/14
26.	#IndianArmy salutes the martyrs who sacrificed their lives for the nation in Mumbai attack on 26 Nov 2008. #26/11.	26/11/15
27.	#MumbaiTerrorAttack Mumbai attack anniversary: US stands with India in its quest for justice, says Trump.	27/11/18
28.	Major Unnikrishnan had lost his life while battling the Pakistani Lashkar-e-Taiba terrorists during the 26/11 attack on Mumbai in 2008.	27/11/18
29.	Army Chief on reports that Pak PM admitted that Mumbai attack was perpetrated by Pak terror group LeT: We know who did it. I don't think we have to get anymore statement from anybody. Int'l community knows who did it. Acceptance is good but even without it,we knew who had done it.	08/12/18
30.	Pakistan Prime Minister Imran Khan had said the 2008 Mumbai attack was *an act of terrorism and* that resolving the case was in Pakistan's interest.	09/12/18

TABLE 4 Fake news dataset 2: Denzel Washington endorses Donald Trump!

Number	News	Date
1.	OMG Denzel Washington endorses Trump/Speaks out against Pres. Obama. #selfhater	03/08/16
2.	Denzel Washington endorses @realDonaldTrump.	03/08/16
3.	ENZEL WASHINGTON ENDORSES DONALD TRUMP!!	03/08/16
4.	Please try to get an interview with Trump Lover Denzel Washington. It offends me that this AA is so dense that he'd endorse a racist.	03/08/16
5.	Good Christian man defies Hollywood endorses Trump!	03/08/16
6.	DENZEL WASHINGTON ENDORSES @realDonaldTrump #MAGA	03/08/16
7.	DENZEL WASHINGTON ENDORSES DONALD TRUMP!!	03/08/16
8.	If the Democrats had won the election, we never would have found out they were using false documents to get warrants to spy on American citizens and political opponents. We never would have known this. Think about it!	05/08/16
9.	Denzel Washington backs Trump in the most epic way possible.	05/08/16
10.	Denzel Washington switches to Trump.	05/08/16
11.	Ice Cube endorses Trump. Liberals crying now. First Denzel Washington, now Ice Cube?	13/08/16
12.	There's more on the Hollywood #TrumpTrain than you'd think.. Now Denzel Washington Endorses Trump shocks Hollywood.	15/09/16
13.	Denzel Washington endorses Donald Trump says Obama murders Christians!	23/09/16
14.	Denzel Washington endorses Trump.	23/09/16
15.	Denzel Washington endorses Donald Trump says Obama murders Christians! -	24/09/16
16.	Oscar winner & Devout Christian Denzel Washington endorses Donald Trump: His devotion to his faith has shocked …	02/10/16
17.	Denzel Washington endorsed Trump, according to AmericaNews, Breitbartt, USANewsHome — and Facebook.	14/10/16
18.	#trumptrain @realDonaldTrump #WednesdayWisdom Denzel Washington endorses @realDonaldTrum.	19/10/16
19.	Denzel Washington endorses Trump! Now that could be a game changer.	21/10/16
20.	Washington Denzel ENDORSES TRUMP! (AND CALLS OBAMA'S AGENDA 'ANTI-CHRISTIAN'!)	22/10/16
21.	Denzel Washington is my favorite actor-movie *déjà vu* now he endorses @realDonaldTrump he = becoming a real live hero.	25/10/16
22.	Denzel Washington endorses Trump we need more jobs unemployment is way up hired more employees more people, than anyone I know in the world.	27/10/16
23.	Denzel Washington explains how his latest character is like a Trump voter – TheBlaze.	28/11/17
24.	Denzel Washington backs Trump in the most epic way possible.	11/01/18
25.	Continue reading Denzel Washington on Trump: 'God puts kings in a place for a season and reason' at Celebitchy.	24/01/18
26.	Denzel Washington: Trump's election saved us from 'Orwellian Police State.'	03/02/18
27.	Denzel Washington: Trump's election saved us from 'Orwellian Police State.'	05/02/18
28.	Denzel Washington: Trump's decision spared us from the hands democrats.	06/02/18
29.	This is the funniest shit I've read since Denzel Washington endorses Donald Trump.	02/08/18
30.	Denzel Washington endorses Trump for his leadership and Franklin Graham and his entire family are supporting Trump. That should tell you something. Read your Bible. Trump is a gift from God.	23/08/18

Note: More datasets on different topics used as part of the research are available upon request from the corresponding author. All the tweets in our datasets were collected before April 25, 2019.

np.random.shuffle (indices_data)
data = data[indices_data]
label = label[indices_data]
politi = politi.sample (frac = 1).reset_index(drop = True)
*nb_test_samples = int(0.4 * data.shape[0])*
x_train = data[:-nb_test_samples]
y_train = labels[:-nb_test_samples]
x_val = data[−nb_test_samples:]
y_val = labels[−nb_test_samples:].

TABLE 5 A fragment of PolitiFact dataset comprising of real news instances

ID	Real News
politifact4358	Democratic leaders say House Democrats are united against GOP Default Act.
politifact8227	Covering young adults under the Affordable Care Act: The importance of outreach and medicaid expansion.
politifact14064	Donald Trump exaggerates when he says China has 'total control' over North Korea.
politifact13132	Hillary Clinton says guns exceed next nine categories as leading cause of death for young black men.
politifact1436	McDonnell letter urges no delay in Virginia offshore energy exploration and development.
politifact13013	UN Refugee Agency welcomes arrival of 10,000th Syrian refugee resettled to United States.
politifact11761	Marco Rubio says Ted Cruz voted for defense cuts in Rand Paul's budget proposal.
politifact8310	No MidOct paycheck for troops if government shuts down lawmaker says.
politifact8130	Florida democrats ask HHS to protect Floridians from high insurance premiums in marketplace.
politifact2734	Amendment to declare English as the national language of the Government of the United States.
politifact6831	Republican consultant says Barack Obama promised to halve federal deficit in his first term.
politifact6652	Obama administration opens the door for states to seek major changes in welfaretowork law.
politifact11552	Leonard Lance claims federal tax code contains 4 million words, is 7 times as long as Bible.
politifact408	Bill Clinton says he 'Never Said A Bad Word About Senator Obama,' At mtvU's First 'Editorial Board.'
politifact2048	Ron Paul praises embattled RNC Chief Michael Steele for leadership On Afghanistan.
politifact6931	The Obameter: Introduce a comprehensive immigration bill in the first year.
politifact4555	Testimony submitted to the Congressional Oversight Panel's hearing on the TARPs impact on financial stability.
politifact11416	Did Bernie Sanders vote against background checks and waiting periods for gun purchases?
politifact440	No regrets for a love of explosives; In a memoir of sorts, a war protester talks of life with the weathermen.
politifact9107	Mitch McConnell, Alison Lundergan Grimes have 'Sharp Differences' on campaign finance rules.

TABLE 6 A fragment of PolitiFact dataset comprising of fake news instances

ID	Fake News
politifact15014	BREAKING: First NFL Team declares bankruptcy over kneeling thugs.
politifact14745	UPDATE: Second Roy Moore accuser works for Michelle Obama right NOW -
politifact14890	Actress Sandra Bullock to Hillary Clinton if You Don't Like Our President You Can Leave and Never Come Back Again You Are One Jealous Woman Who Ls Nothing to Compare With Trump I Hope He Will Arrest Y
politifact15052	Inside the Snapchat bloodbath: 60 minutes of terror which left 17 dead, bodies in classrooms and kids 'full of blood.'
politifact14388	NYC: PHYSICIAN and wife jump to death¦Leave kids behind because they can't afford the health care only hours after McCain, Murkowski, Collins and EVERY democrat voted to keep Obamacare -
politifact13677	BREAKING: Meryl Streep just got fired from a major project for lying about Trump -
politifact15130	Could Trump win the Nobel peace prize? Peace in Korean Peninsula would be significant foreign policy achievement.
politifact14694	UPDATE: Hillary Clinton leaves the country as Mueller Indictment is announced.
politifact13905	SPECIAL REPORT: GEORGIA BECOMES FIRST STATE TO BAN MUSLIM CULTURE IN HISTORIC MOVE TO RESTORE WESTERN VALUES!
politifact14126	Breaking News: FBI uncovers evidence that 62 million Trump voters are all Russian agents.
politifact14155	BOMBSHELL: COMEY KNEW MURDERED DNC STAFFER, SETH RICH, WAS WIKILEAKS SOURCE & COVERED IT UP FOR HILLARY.
politifact14904	HOLLYWOOD CELEBS: WE WILL GO ON A TOTAL STRIKE IF TRUMP DOES NOT RESIGN.
politifact13544	BREAKING: Florida moves for FULL RECOUNT of state over massive voter fraud (DETAILS).
politifact14960	Haiti official getting ready to testify against Clinton Foundation corruption NEXT WEEK, found DEAD with a gunshot to the head.
politifact15100	Kushner And wife Ivanka Trump were tossed out of the White House and Donald is cutting them out of his will!
politifact13934	BREAKING: KEN STARR'S PLANE JUST DISAPPEARED ON HIS WAY TO DC TO TESTIFY AGAINST HILLARY.
politifact14145	DONALD TRUMP PRAISES COLONEL SANDERS FOR HIS SERVICE IN THE CIVIL WAR Manny Schewitz O 2 days ago off the wire speaking to our intern Dave Robicheaux who is paid in Facebook likes Trump insisted Col H
politifact14818	FBI uncovered Russian bribery plot before Obama administration approved controversial nuclear deal with Moscow.
politifact14026	Fox News Sandra Smith: Everybody could see from the United Airlines passenger's behavior that he was not a genuine American.
politifact14490	BREAKING NEWS: Prince William and Harry donates $100 million to Hurricane Harvey victims.

Fourth, we use GloVe (Global Vectors used for retrieving word vector representation)[20] that includes 6 billion tokens each of 100 dimensions to perform embedding and create an embedded matrix. Embedding converts each word in the sequence to n-dimensional vectors where n is the embedding dimension.

```
embedding_index = {}
file = open('glove.6B.100d.txt',encoding = 'utf8')
 for line in file:
    value = line.split()
    word = value[0]
    coefs = np.asarray (value[1:], dtype = 'float32')
    embedding_index[word] = coefs
file.close().
```

Fifth, the resultant embedding matrix is used to create an embedding layer. The output of this embedding layer is passed onto the further layers of our model.

```
embedding_matrix = np.random.random((len (word_indexes) + 1, 100))
for word, i in word_indexes.items():
    embedding_vector = embedding_index.get (word)
    if embedding_vector is not None:
        embedding_matrix[i] = embedding_vector
embeddings_layer = Embedding (len (word_indexes) + 1,100,weights = [embedding_matrix],
input_length = 1000,trainable = True).
```

3.3 | Model preparation

For sentiment classification, we carefully select 7 models, which incorporate variations of the convolutional neural network (CNN) and long short-term memory (LSTM) architectures as part of our analysis. The CNN models are widely used for image classification and detection purposes while also being used for sentence classification.[21] Simple RNNs were not employed in our case due to restriction in their information retention capability and because of vanishing and exploding gradient problems.[22] That is why, to filter out these issues with RNNs, we made use of LSTMs and its modification, the bidirectional LSTMs. For pooling in our network, MaxPooling is used, which is the most common pooling technique. It is done by applying a *max filter* to (usually) nonoverlapping subregions of the initial representation. In addition, to map the results, we use the Leaky ReLU activation function instead of rectified linear unit (ReLU) because when using the ReLU activation function, negative values becoming zero immediately decreased the accuracy of our model as well as its ability to fit or train from the data properly.

Moving ahead of simpler models, we implement their combinations using ensembling. The method of ensembling in various networks has proven to be highly efficient in increasing the performance of a network.[23,24] It involves combining different models, which when implemented individually do not perform well. Different types of ensembling methods are bagging, boosting, and stacking. Another factor we use is the attention mechanism. In recent years, attention has emerged out as a widely used and important tool in field of deep learning.[25-28] Attention can be defined as a vector derived as output of dense layer of network using the softmax function. Before attention, one needed to compress all the input information into a fixed length vector. However, a sentence with large number of words would result in loss of important information. Attention partially fixes this problem by helping the network in learning as to where it should pay attention in the input sequence for each item in the output sequence.

4 | MODEL ANALYSIS

This section covers the implementation details about our models on the PolitiFact dataset having 634 training samples and 422 test samples and is structured into 7 parts:

i. CNN model;
ii. LSTM model;
iii. Bidirectional LSTM model;

iv. CNN + LSTM ensembled model;
v. Bidirectional LSTM + LSTM ensembled model;
vi. CNN + LSTM ensembled model with attention mechanism;
vii. CNN + bidirectional LSTM ensembled model with attention mechanism.

Additionally, comparison with machine learning algorithms has also been done in the continued subsection to measure their relative performance with deep learning methods on the same datasets. These algorithms include logistic regression and SVM.

For all the abovementioned models, we use *categorical cross entropy* loss as a loss function. After Bayesian optimization method for hyperparameter tuning, we obtain the set of optimal hyperparameters. GPy and GPyOpt libraries are used for implementation in Python. The domain of hyperparameters that we use is as follows:

Learning rate = [0.1, 0.01, 0.001],
Optimizers = [SGD, Adam, RMSprop],
Epochs =[15, 20],
Batch size = [32, 64, 128].

Due to resource limitations, our team could not try out the same procedure for complex models such as ensemble networks or the models implementing attention mechanism. Therefore, on the basis of evaluation on simpler models, we set the hyperparameters for the complex models accordingly. Further work can be done by trying out the abovementioned approach on all the models and also by extending the domain.

In our models, the learning rate is set to 10e-3 and the training process is executed for 15 epochs with RMSprop optimizer on the training samples with batch size set to 128 and using the test set for the calculation of accuracy. We use *ModelCheckpoint callback* as a way of monitoring the accuracy and saving the weights of the model with best accuracy. The test set accuracy is computed by comparing the predictions of the model on the test set with the actual labels. The accuracy is then formulated as the ratio of correct predictions to the total predictions and it was then multiplied by 100. Using the above procedure, we calculate 3 accuracies for different randomly shuffled datasets and compute the average of these 3 accuracies for our comparison.

```
bounds = [{‘name’:’optimizer’,’type’:’discrete’,’domain’:(0,1,2)},
{‘name’:’lr’,’type’:’discrete’,’domain’:(0.001,0.01,0.1)},
{‘name’:’batch_size’,’type’:’discrete’,’domain’:(32,64,128)},
{‘name’:’epochs’,’type’:’discrete’,’domain’:(15, 20)}]

opt_model = GPyOpt.methods. BayesianOptimization(f = f, domain = bounds),
opt_model.run_optimization(max_iter = 10),
```

where *f* is a function to train the model with the hyperparameters specified in the bounds array.

4.1 | CNN model

In the first phase of our comparative study, we apply simple CNN with 3 convolution blocks, each consisting of a single Conv-1D and a MaxPooling layer. In addition to these blocks, a flatten layer and a fully connected layer with 128 nodes were added. Finally, this model classifies fake and real news using LeakyReLU activation function (Figure 2). The average accuracy was 73.29%.

```
seq_input = Input (shape = (1000,), dtype = ‘int32’)
embedded_sequences = embeddings_layer(seq_input)
cov1 = Conv1D(128,5,activation = ‘relu’)(embedded_sequences)
pool1 = MaxPooling1D(5)(cov1)
cov2 = Conv1D(128, 5, activation = ‘relu’)(pool1)
pool2 = MaxPooling1D(5)(cov2)
cov3 = Conv1D(128, 5, activation = ‘relu’)(pool2)
pool3 = MaxPooling1D(35)(cov3)
flat = Flatten()(pool3)
```

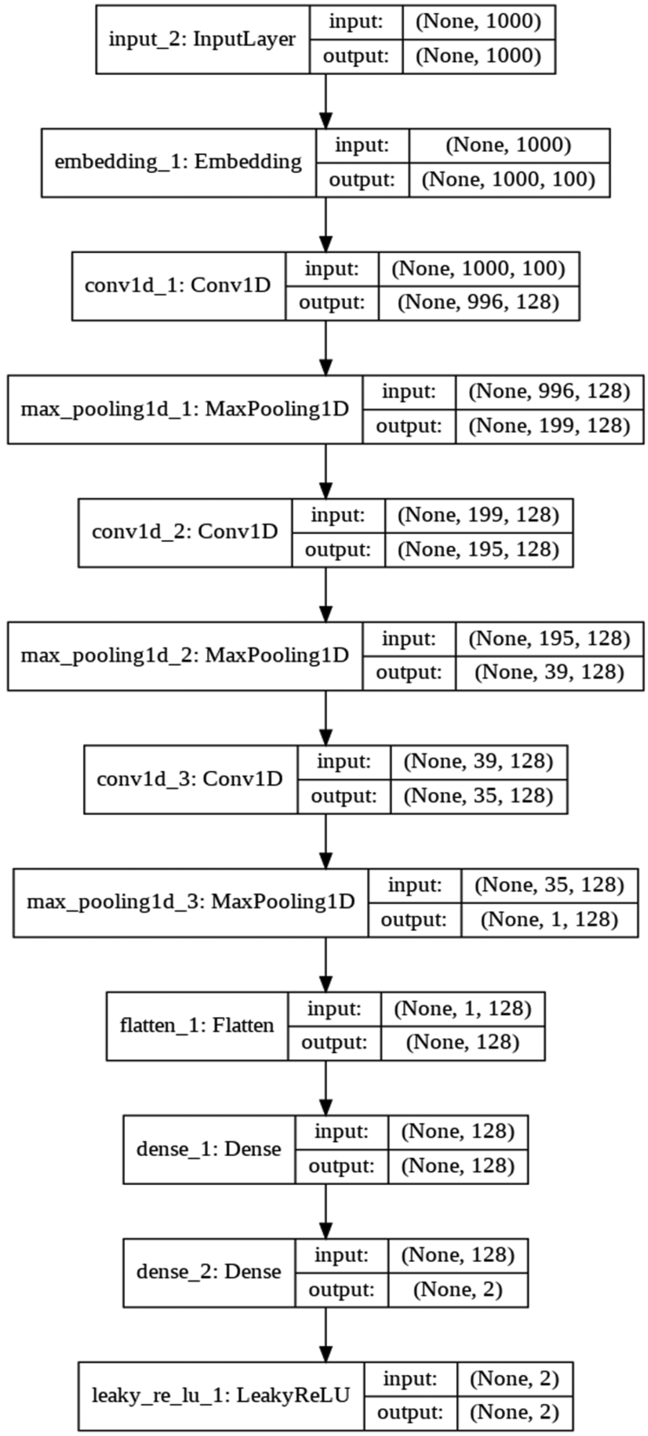

FIGURE 2 Convolutional neural network (CNN) model

dense = Dense(128, activation = 'relu')(flat)

preds = Dense (len (macronum))(dense)
reLU = LeakyReLU (alpha = 0.1)(preds)

model = Model (seq_input, reLU).

4.2 | LSTM model

We then try the LSTM model as LSTMs are successful in overcoming the issue of long-term dependencies (one of the major limitations of traditional RNN) with their property of remembering information for long duration of time.

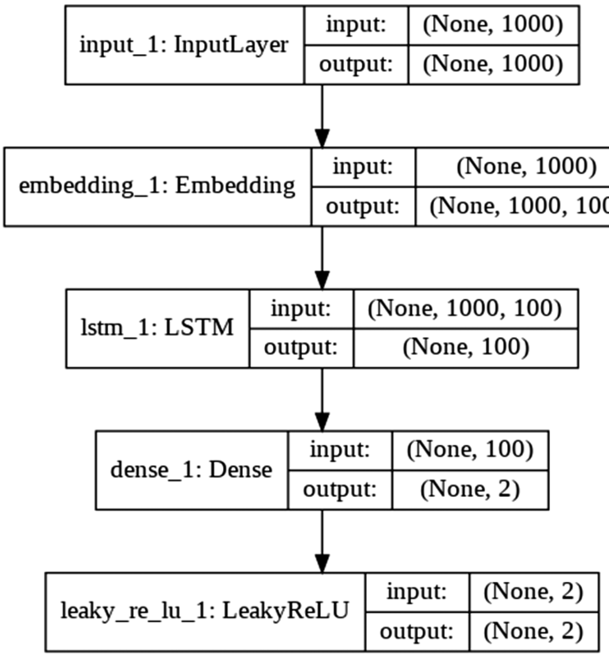

FIGURE 3 Long short-term memory (LSTM) model

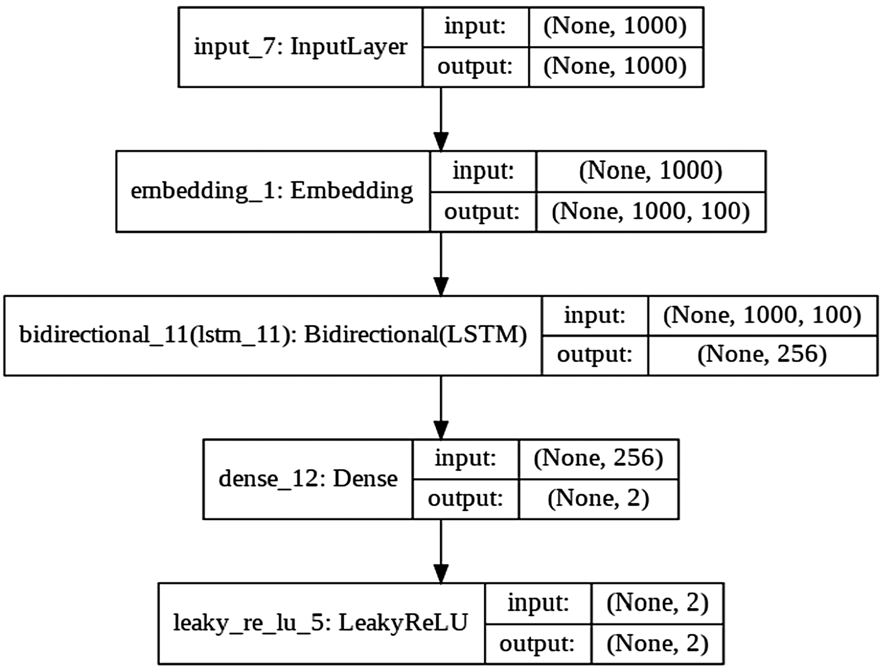

FIGURE 4 Bidirectional long short-term memory (LSTM) model

The embedding layer's output is fed into the LSTM unit for computation (Figure 3). The resulting (average) accuracy of this model was 80.62%, which is a significant improvement over our CNN model and proves that LSTM architecture is much more powerful for classification tasks.

seq_input = Input (shape = (1000,), dtype = 'int32')
embedded_sequences = embeddings_layer(seq_input)
lstm = LSTM(100)(embedded_sequences)

preds = Dense (len (macronum))(lstm)
reLU = LeakyReLU (alpha = 0.1)(preds)

model = Model (seq_input, reLU).

4.3 | Bidirectional LSTM model

In this network, we use the bidirectional LSTM model, which can be defined as an extension of traditional LSTM and is used to increase model efficiency for various NLP problems. Bidirectional LSTM trains two LSTMs instead of one (as in case of traditional LSTM network). It is widely used in various popular problems including speech recognition, translation, handwriting recognition, etc.

The dense layer receives the output from the bidirectional LSTM unit to produce the output (Figure 4). This network architecture gave an average accuracy of 83.81%.

```
seq_input = Input (shape = (1000,), dtype = 'int32')
embedded_sequences = embeddings_layer(seq_input)
bi2 = Bidirectional (LSTM(128))(embedded_sequences)

preds = Dense (len (macronum))(bi2)
reLU = LeakyReLU (alpha = 0.1)(preds)

model = Model (seq_input, reLU).
```

4.4 | CNN + LSTM ensembled model

Moving further, we try ensembling the CNN and LSTM models. Ensembling can be used on both supervised learning tasks as well as for unsupervised learning. Our model uses the predictions of a few basic classifiers as a base (first level), and then uses another model at the second level to predict the output from the earlier predictions at the base level. We combine the CNN and LSTM architectures by using 3 Conv-1D and 3 MaxPooling layers along with the LSTM layer (Figure 5). The results were comparable to the previous two models as this network achieved an average accuracy of 86.14%.

```
seq_input = Input (shape = (1000,), dtype = 'int32')
embedded_sequences = embeddings_layer(seq_input)
cov1 = Conv1D(128, 5)(embedded_sequences)
pool1 = MaxPooling1D(5)(cov1)
cov2 = Conv1D(128, 5)(pool1)
pool2 = MaxPooling1D(5)(cov2)
cov3 = Conv1D(128, 5)(pool2)
pool3 = MaxPooling1D(35)(cov3)
flat = Flatten()(pool3)
dense = Dense(128)(flat)
zeros = Lambda (lambda x: K.zeros_like(x), output_shape = lambda s: s)(dense)

rnn_layer = LSTM(128, return_sequences = False, batch_input_shape = (10, 1000,), stateful = False)(embedded_sequences, initial_state = [dense, dense])
preds = Dense (len (macronum), activation = 'softmax')(rnn_layer)
reLU = LeakyReLU (alpha = 0.1)(preds)

model = Model (seq_input, reLU).
```

4.5 | Bidirectional LSTM + LSTM ensembled model

This approach combines both the LSTM as well as the Bidirectional LSTM using ensembling and the concatenated unit is then fed to the dense layer (Figure 6). This ensembled model delivered an increased (average) accuracy of 86.89%, partly due to the varying combination of LSTMs.

```
seq_input = Input (shape = (1000,), dtype = 'int32')
embedded_sequences = embeddings_layer(seq_input)
```

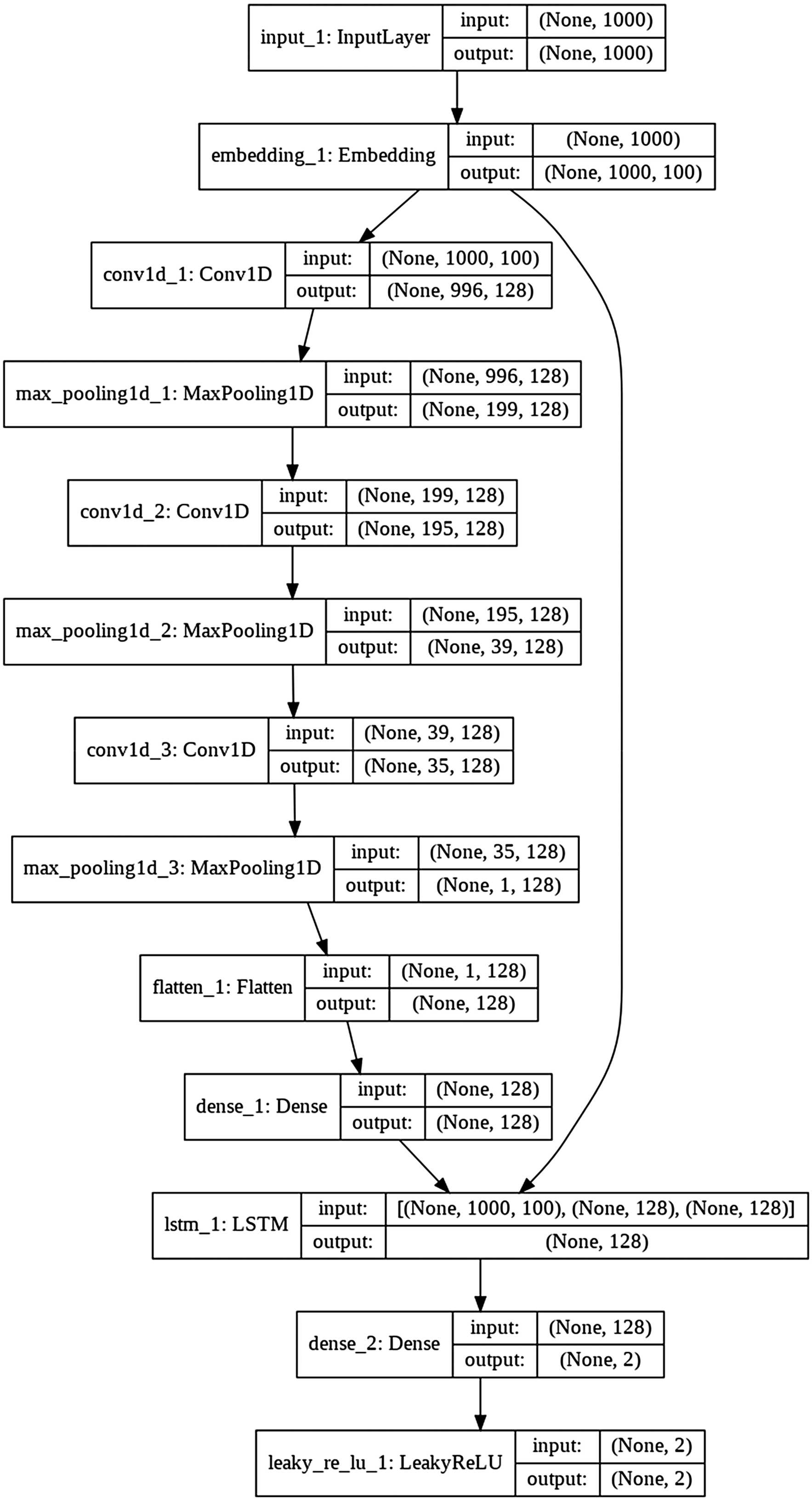

FIGURE 5 CNN + LSTM ensembled model. CNN, convolutional neural network; LSTM, long short-term memory

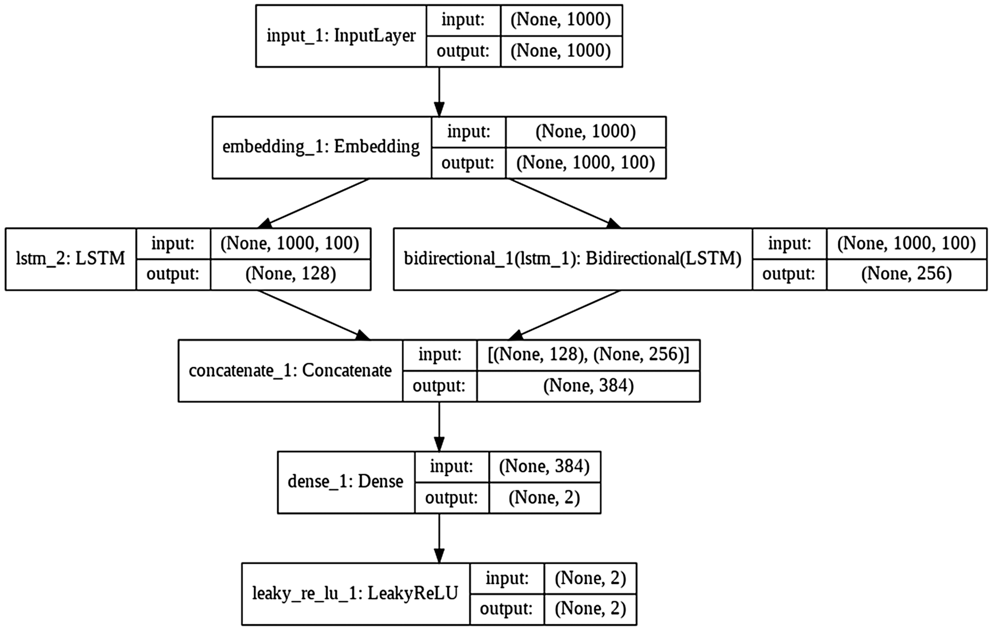

FIGURE 6 Bidirectional LSTM + LSTM ensembled model. LSTM, long short-term memory

bi2 = Bidirectional (LSTM(128))(embedded_sequences)
bi1 = (LSTM(128))(embedded_sequences)
dense = concatenate([bi1,bi2])

preds = Dense (len (macronum))(dense)
leakyReLU = LeakyReLU (alpha = 0.1)(preds)

model = Model (seq_input, leakyReLU).

4.6 | CNN + LSTM ensembled model with attention mechanism

In this model, we combine the CNN and the traditional LSTM architectures. Like the similar CNN + LSTM model, this model consists of CNN with 3 convolution layers; however, in this approach, we also employ the use of attention layer that would help the model to learn to pay attention only to specific areas of input sequences instead of working on the entire series of input sequences (Figure 7).

Using attention mechanism proved useful by a slight margin as this one gave a better average accuracy of 86.57%.

seq_input = Input (shape = (1000,), dtype = 'int32')
embedded_sequences = embeddings_layer(seq_input)
cov1 = Conv1D(128, 5)(embedded_sequences)
pool1 = MaxPooling1D(5)(cov1)
cov2 = Conv1D(128, 5)(pool1)
pool2 = MaxPooling1D(5)(cov2)
cov3 = Conv1D(128, 5)(pool2)
pool3 = MaxPooling1D(35)(cov3) # global max pooling
flat = Flatten()(pool3)
dense = Dense(128)(flat)
zeros = Lambda (lambda x: K.zeros_like(x), output_shape = lambda s: s)(dense)
rnn_layer = LSTM(128,return_sequences = False, batch_input_shape = (10,1000,),
stateful = False)(embedded_sequences,initial_state = [dense, dense])
attention_prob = Dense(128,activation = 'softmax', name = 'attention_vec')(rnn_layer)

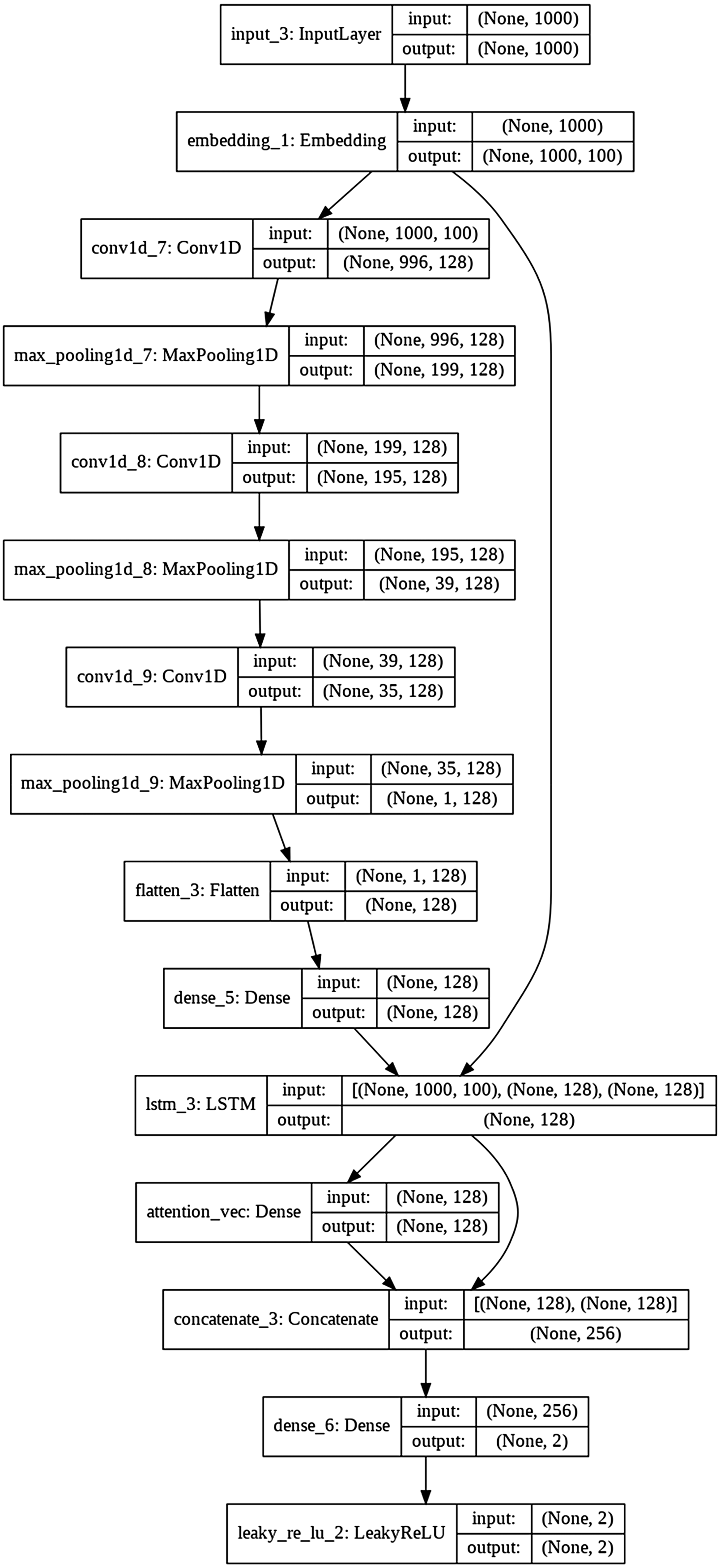

FIGURE 7 CNN + LSTM ensembled model with attention mechanism. CNN, convolutional neural network; LSTM, long short-term memory

```
attention_mul = concatenate([rnn_layer, attention_prob])
preds = Dense (len (macronum),activation = 'softmax')(attention_mul)
leaky = LeakyReLU (alpha = 0.1)(preds)

model = Model (seq_input, leaky).
```

4.7 | CNN + bidirectional LSTM ensembled model with attention mechanism

Our final model consists of CNN with 3 convolution layers and Leaky ReLU activation function combined with bidirectional LSTM network for better contextual understanding (Figure 8). In this model, we have again used attention mechanism to increase our efficiency as well as accuracy. As a result, this model, on average, achieved state-of-the-art accuracy of 88.78%, the highest in our study.

```
seq_input = Input (shape = (1000,), dtype = 'int32')
embedded_sequences = embeddings_layer(seq_input)
cov1 = Conv1D(128,5,activation = 'relu')(embedded_sequences)
pool1 = MaxPooling1D(5)(cov1)
cov2 = Conv1D(128, 5, activation = 'relu')(pool1)
pool2 = MaxPooling1D(5)(cov2)
cov3 = Conv1D(128, 5, activation = 'relu')(pool2)
pool3 = MaxPooling1D(35)(cov3)
flat = Flatten()(pool3)
dense = Dense(128, activation = 'relu')(flat)
zeros = Lambda (lambda x: K.zeros_like(x), output_shape = lambda s: s)(dense)
bi2 = Bidirectional (LSTM(128))(embedded_sequences)
denselayer = concatenate([dense,bi2])
attention_prob = Dense(128,activation = 'softmax', name = 'attention_vec')(denselayer)
attention_mul = concatenate([denselayer, attention_prob])

preds = Dense (len (macronum))(attention_mul)
leaky = LeakyReLU (alpha = 0.1)(preds)

model = Model (seq_input, leaky).
```

4.8 | Comparison with machine learning algorithms

To improve the significance of the results, we compare our results with classifiers such as logistic regression and SVM. In both of these classifiers, we do not convert the labels into categorical arrays.

4.8.1 | Logistic regression

First, we train logistic regression algorithm on the preprocessed training set of both the datasets. Next, we predict the test samples and compute the accuracies.

```
from sklearn.linear_model import LogisticRegression
logistic = LogisticRegression (random_state = 0,solver = 'lbfgs',multi_class = 'multinomial').fit(x_train,y_train)
y_preds = logistic.predict(x_val)

from sklearn.metrics import accuracy_score
accuracy_score(y_val,y_preds).
```

The average accuracy in this case was 57.58%.

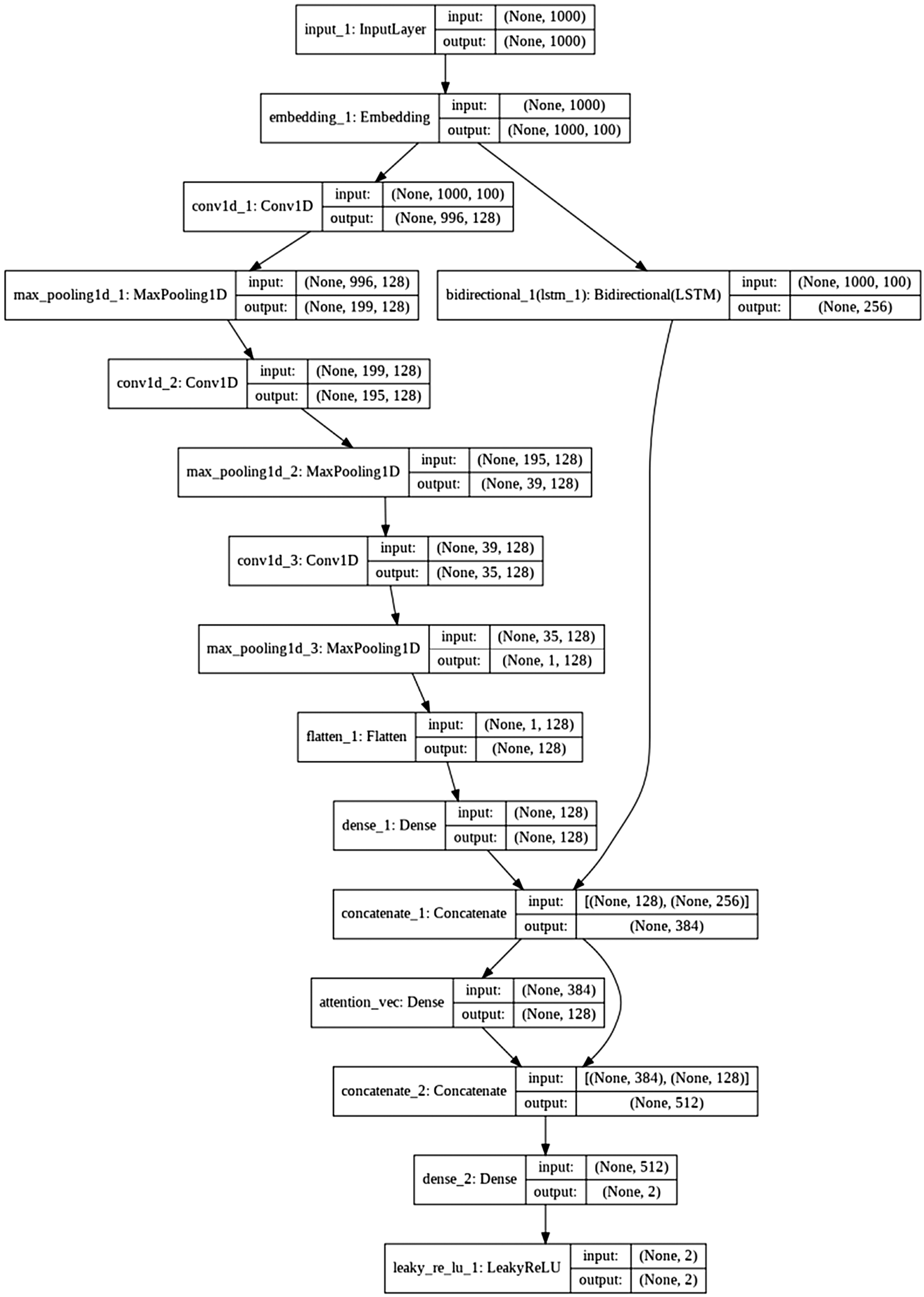

FIGURE 8 CNN + bidirectional LSTM ensembled model with attention mechanism. CNN, convolutional neural network; LSTM, long short-term memory

4.8.2 | Support vector machine

We then use SVM as our next baseline for comparison with our models and the average accuracy obtained here was 58.68%.

```
from sklearn.svm import SVC
supportvector = SVC().fit(x_train,y_train)
y_preds = supportvector.predict(x_val)

from sklearn.metrics import accuracy_score
accuracy_score(y_val,y_preds).
```

5 | RESULTS

Our created datasets resulted in testing accuracy of ~97% in almost every model due to its small size. Hence, we tested our model on a much larger dataset, ie, PolitiFact. In terms of the accuracies achieved by the models, we found the following averaged results on the PolitiFact dataset:

i. 73.29% in the CNN model;
ii. 80.62% in the LSTM model;
iii. 83.81% in the bidirectional LSTM model;
iv. 86.14% in the CNN + LSTM ensembled model;
v. 86.89% in the bidirectional LSTM + LSTM ensembled model;
vi. 86.57% in the CNN + LSTM ensembled model with attention mechanism;
vii. 88.78% in the CNN + Bidirectional LSTM ensembled model with attention mechanism.

For the machine learning algorithms, the results were as follows:

i. 57.58% in logistic regression;
ii. 58.68% in SVM.

These results prove that ensemble networks perform comparatively better when compared to simpler architectures. Moreover, attention mechanism, when used with the ensemble network of CNN and bidirectional LSTM, showed a remarkable accuracy of 88.78%, the highest accuracy achieved in our study.

In order to ensure that the difference in their mean accuracies is statistically significant, we use null hypothesis statistical testing. Null hypothesis states that the performance metric of both the models is equal and the small difference between their mean accuracies is statistically insignificant. For each pair of model, we calculate a test statistic and a p-value using paired sample t-test.

For each pair of model, we create an array that consists of difference of accuracies of both the models on each split of dataset. Then, we calculate the mean of the set of values in the difference array. Next, we calculate the standard deviation of the set of values in the difference array and calculate the test statistic using mean and standard deviation. Finally, p-value is obtained by comparing t-statistic to a t-distribution with (n-1) degrees of freedom, where n is the number of splits of dataset, which in our case was 3. If p-value is below the threshold, ie, 0.05, we reject the null hypothesis and infer that the difference is statistically significant.

The p-values calculated for different pair of models are shown in Figure 9.

6 | DISCUSSION

Many previous methods have focussed their research on social bots and their effect toward the spread of false information. Our method, instead, focusses on the differentiating factor of fake and real news stories, ie, their sentiment. For this research, our team collected and prepared the required data from news sources such as Twitter and PolitiFact. These datasets contained both real as well as the fake news stories. For fact-checked news, we used the dataset from PolitiFact, which publishes the original statement of political news articles and their complete fact-checked evaluation results to

Models	CNN	LSTM	Bidirectional LSTM	CNN+LSTM	Bidirectional LSTM+LSTM	CNN+LSTM with attention	CNN+bidirectional LSTM with attention
CNN		.018477	. 017552	.007456	.008502	.002392	.004888
LSTM	.018477		.254526	.04508	.02342	.006096	.011603
Bidirectional LSTM	.017552	.254526		.674463	.208724	.308251	.192122
CNN+LSTM	.007456	.04508	.674463		.515385	.391438	.001656
Bidirectional LSTM+LSTM	.008502	.02342	.208724	.515385		.723433	.319409
CNN+LSTM with attention	.002392	.006096	.308251	.391438	.723433		.110473
CNN+bidirectional LSTM with attention	.004888	.011603	.192122	.001656	.319409	.110473	

FIGURE 9 p-values for different pair of models. CNN, convolutional neural network; LSTM, long short-term memory

mark them as fake or real. After data collection, we performed in-depth analysis of the data and found the differing patterns in both these datasets. The fake news dataset carried emotions or sentiments of an *exciting nature*, whereas the real news dataset contained monotonous sentiments. Moreover, characters such as "!", "?" often appeared randomly in the sentences and unusual capitalization was also observed. All of these helped us to create a hypothesis that the sentiments of news stories would play a major role toward the solution of our problem.

Deep neural networks are widely used for sentiment analysis and thus were an efficient solution to our problem.[29-32] Firstly, we extracted each news title from the datasets using BeautifulSoup library in Python. Secondly, after splitting each sequence into words, we retrieved word vector representations and embedded the data that we collected using GloVe. The resulting embedded matrix is used to create embedding layer in which the input was fed.

Initially, we tested several simple models such as CNNs, LSTMs, and bidirectional LSTMs. The results were satisfactory but not promising. The CNN architecture gave the lowest accuracy in comparison to the others that we studied. The LSTM architecture and bidirectional LSTM architecture performed significantly better in comparison to simple CNN architecture. We further increased our appetite for improved accuracy and incorporated more complex models as part of our methodology.

Such models included advanced techniques, ie, ensembling and attention mechanism. Next, we trained the data on an ensembled CNN and LSTM architecture as well as the ensemble network of bidirectional LSTM and LSTM. We observed that there was more increase in the accuracy when bidirectional LSTM was used in comparison to CNN. Next, we incorporated the attention mechanism into our last two models and formed the ensemble model of CNN and LSTM as well as the ensemble network architecture of CNN and bidirectional LSTM with attention mechanism. Finally, we concluded that the CNN + bidirectional LSTM ensembled model with attention mechanism performed the best among all the architectures that were studied in this research.

However, in a real-life scenario, there exist more subtle forms of fake news such as incorrectly reported facts by a trustable source and transition of real news to fake news during its propagation. Although, deep learning methodology is quite powerful for complex problems, fact-checking for real-life events still remains an unsolved problem in the domain of deep learning and this should be done manually. Many websites such as PolitiFact and FactCheck.org perform trustable fact checking, which can help news sources from getting exploited. The second issue can be treated by regularly checking the news stories shared by the public with the actual source and obtaining the semantic difference between them. If the difference is greater than a specified threshold, we can classify them as real news transformed into fake news.

7 | CONCLUSION

We can use this research to combat fake news stories and neutralize the drastic effects of false information on a large scale. Our study can also be employed as a framework and an effective tool for in-depth survey on this topic. Moving further, educational efforts regarding digital media can help in informing the public in the social media about the pressing issues and thus minimize the misperceptions surrounding them.[33]

Finally, we would like to address the shortcomings of this study. First, this research focussed mostly on the sentiments of news stories while not paying continuous attention to the credibility of the news sources themselves, due to resource limitations. An even greater accuracy can be achieved by keeping track of the producer (source) of the fake news story and then including it as a parameter in the research. Second, our classification models were unable to identify the semantic transition of real news to fake news during its propagation mainly due to the comparative nature of our study. Moreover, this research may be expanded to achieve better accuracy by using other newly developed state-of-the-art architectures.

ORCID

Sachin Kumar https://orcid.org/0000-0002-5324-2156

REFERENCES

1. Friggeri A, Adamic L, Eckles D, Cheng J. Rumor cascades. Paper presented at: Eighth International AAAI Conference on Weblogs and Social Media; 2014; Ann Arbor, MI.
2. Allcott H, Gentzkow M. Social media and fake news in the 2016 election. *J Econ Perspect*. 2017;31(2):211e236. https://doi.org/10.1257/jep.31.2.211
3. Donald B. Stanford researchers find students have trouble judging the credibility of information online. https://ed.stanford.edu/news/stanford-researchers-find-students-have-trouble-judging-credibility-information-online. 2016.
4. Journalism, Fake News & Disinformation. https://en.unesco.org/sites/default/files/journalism_fake_news_disinformation_print_friendly_0.pdf. 2018.
5. How is Fake News Spread? Bots, People like You, Trolls, and Microtargeting. http://www.cits.ucsb.edu/fake-news/spread. 2018.
6. Lohr S. It's True: False News Spreads Faster and Wider. And Humans Are to Blame. https://www.nytimes.com/2018/03/08/technology/twitter-fake-news-research.html. 2018.
7. Lyons T. Hard Questions: What's Facebook's Strategy for Stopping False News? Facebook. 2018. https://newsroom.fb.com/news/2018/05/hard-questions-false-news/. Accessed May 23, 2018.
8. Shao C, Ciampaglia GL, Varol O, Flammini A, Menczer F. The spread of low-credibility content by social bots. ArXiv:1707.07592v4 [cs.SI]. 2018.
9. Ferrara E, Varol O, Davis C, Menczer F, Flammini A. The rise of social bots. In: Proceedings of the 25th International Conference Companion on World Wide Web; 2016; Montreal, Canada.
10. Zhao Z, Resnick P, Mei Q. Enquiring minds: early detection of rumors in social media from enquiry posts. In: Proceedings of the Twenty-Fourth International Conference World Wide Web; 2015; Florence, Italy.
11. Shelke S, Attar V. Source detection of rumor in social network – a review. *Online Soc Netw Media*. 2019;9:30-42. https://doi.org/10.1016/j.osnem.2018.12.001
12. Gilda S. Evaluating machine learning algorithms for fake news detection. Paper presented at: 2017 IEEE 15th Student Conference on Research and Development (SCOReD); 2017; Putrajaya, Malaysia.
13. Ruchansky N, Seo S, Liu Y. CSI: a hybrid deep model for fake news detection. In: Proceedings of the 2017 ACM on Conference on Information and Knowledge Management (CIKM '17); 2017; Singapore.
14. Granik M, Mesyura V. Fake news detection using naive Bayes classifier. Paper presented at: 2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON); 2017; Kiev, Ukraine.
15. Ko H, Hong JY, Kim S, Mesicek L, Na IS. Human-machine interaction: a case study on fake news detection using a backtracking based on a cognitive system. *Cogn Syst Res*. 2019;55:77-81. https://doi.org/10.1016/j.cogsys.2018.12.018
16. Shu K, Mahudeswaran D, Wang S, Lee D, Liu H. FakeNewsNet: a data repository with news content, social context and spatial temporal information for studying fake news on social media. ArXiv:1809.01286v3 [cs.SI]. 2019.
17. Shu K, Sliva A, Wang S, Tang J, Liu H. Fake news detection on social media: a data mining perspective. *SIGKDD*. 2017;19(1):22-36. https://doi.org/10.1145/3137597.3137600
18. Shu K, Wang S, Liu H. Beyond news contents: the role of social context for fake news detection. ArXiv: 1712.07709v2 [cs.SI]. 2018.
19. Python Software Foundation. Python Library Reference: beautifulsoup4 v4.7.1. https://pypi.org/project/beautifulsoup4/
20. Pennington J, Socher R, Manning CD. GloVe: global vectors for word representation. In: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP); 2014; Doha, Qatar.
21. Kim Y. Convolutional neural networks for sentence classification. ArXiv: 1408.5882v2 [cs.CL]. 2014.
22. Lipton ZC, Berkowitz J, Elkan C. A critical review of recurrent neural networks for sequence learning. ArXiv: 1506.00019v4 [cs. LG]. 2015.

23. Fawaz HI, Forestier G, Weber J, Idoumghar L, Muller PA. Deep neural network ensembles for time series classification. ArXiv: 1903.06602 [cs. LG]. 2019.
24. Tao S. Deep neural network ensembles. ArXiv: 1904.05488 [cs. LG]. 2019.
25. Bengio Y. The consciousness prior. ArXiv: 1709.08568 [cs. LG]. 2017.
26. Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need. ArXiv: 1706.03762v5 [cs.CL]. 2017.
27. Wang C, Han D, Liu Q, Luo S. A deep learning approach for credit scoring of peer-to-peer lending using attention mechanism LSTM. *IEEE Access*. 2018;7:2161-2168. https://doi.org/10.1109/ACCESS.2018.2887138
28. Young T, Hazarika D, Poria S, Cambria E. Recent trends in deep learning based natural language processing. ArXiv: 1708.02709v8 [cs.CL]. 2018.
29. Zhang L, Wang S, Liu B. Deep learning for sentiment analysis: a survey. ArXiv: 1801.07883v2 [cs.CL]. 2018.
30. Stojanovski D, Strezoski G, Madjarov G, Dimitrovski I, Chorbev I. Deep neural network architecture for sentiment analysis and emotion identification of Twitter messages. *Multimed Tools Appl*. 2018;77:32213-32242. https://doi.org/10.1007/s11042-018-6168-1
31. Hassan A, Mahmood A. Deep learning approach for sentiment analysis of short texts. Paper presented at: 2017 3rd International Conference on Control, Automation and Robotics (ICCAR); 2017; Nagoya, Japan.
32. Severyn A, Moschitti A. Twitter sentiment analysis with deep convolutional neural networks. In: Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '15); 2015; Santiago, Chile.
33. Mele N, Lazer D, Baum M, et al. Combating Fake News: An Agenda for Research and Action. https://shorensteincenter.org/wp-content/uploads/2017/05/Combating-Fake-News-Agenda-for-Research-1.pdf. 2017.

How to cite this article: Kumar S, Asthana R, Upadhyay S, Upreti N, Akbar M. Fake news detection using deep learning models: A novel approach. *Trans Emerging Tel Tech*. 2020;31:e3767. https://doi.org/10.1002/ett.3767