

Article

TChecker: A Content Enrichment Approach for Fake News Detection on Social Media

Nada GabAllah *, Hossam Sharara  and Ahmed Rafea 

Computer Science and Engineering Department, The American University in Cairo, Cairo 11835, Egypt; hossam.sharara@aucegypt.edu (H.S.); rafea@aucegypt.edu (A.R.)

* Correspondence: nadaaym@aucegypt.edu

Abstract: The spread of fake news on social media continues to be one of the main challenges facing internet users, prohibiting them from discerning authentic from fabricated pieces of information. Hence, identifying the veracity of the content in social posts becomes an important challenge, especially with more people continuing to use social media as their main channel for news consumption. Although a number of machine learning models were proposed in the literature to tackle this challenge, the majority rely on the textual content of the post to identify its veracity, which poses a limitation to the performance of such models, especially on platforms where the content of the users' post is limited (e.g., Twitter, where each post is limited to 140 characters). In this paper, we propose a deep-learning approach for tackling the fake news detection problem that incorporates the content of both the social post and the associated news article as well as the context of the social post, coined TChecker. Throughout the experiments, we use the benchmark dataset FakeNewsNet to illustrate that our proposed model (TChecker) is able to achieve higher performance across all metrics against a number of baseline models that utilize the social content only as well as models combining both social and news content.

Keywords: fake news; BERT; BERTweet; BiLSTM; social media



Citation: GabAllah, N.; Sharara, H.; Rafea, A. TChecker: A Content Enrichment Approach for Fake News Detection on Social Media. *Appl. Sci.* **2023**, *13*, 13070. <https://doi.org/10.3390/app132413070>

Academic Editors: Jae-Hoon Kim and Kichun Lee

Received: 10 October 2023
Revised: 28 November 2023
Accepted: 30 November 2023
Published: 7 December 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In the era of Web 2.0, our interactions as well as our perceptions of information are changing. The ways of communication have evolved in recent decades in very fast and ground-breaking ways that are pushing some of the traditional ways for acquiring information into obsolescence. One of the most observable changes is the media, where in the near past television, radio, and newspapers were the primary credible sources of news and information for everyone. Reporters used to race to the locations of events to get the first scope. Live TV coverage of important events could get millions of people watching their TVs at the same time. Headlines in newspapers could change stock market values. Nowadays, it is very rare to find someone reading a newspaper, or watching news on TV; instead, social media is becoming the most prominent source for obtaining information about a news event.

Extracting information from social media has become a very rich area of research, as it has become one of the fastest growing sources of data about almost everything. As Web 2.0 and social media enabled internet users to contribute through their use, this allowed anyone to post and share data about anything, which in turn created a huge repository of data for everyone to access. However, with the power given to everyone to post and share anything on social media comes great responsibility for the content being posted or shared. Unfortunately, social media users do not usually tend to validate or fact-check their posts before sharing them, as they tend to believe what is shared many times within their circle of friends on social media. This phenomenon has been studied in the literature and is known as the validity effect [1], where people tend to have more belief in what is shared

through their close circles as this emphasizes their feeling of validation. In addition, users will tend to share more posts that are aligned with their ideas and previous knowledge, regardless of the truthfulness of these ideas, which is known as confirmation bias [2]. These combined phenomena gives individuals a false sense of credibility about any piece of news that is shared in their circle of acquaintance, and thus share it themselves to other circles, etc., thus leading to less credibility of the news sources themselves, which most often rely on social media as well to gather information [3].

A study presented in [4] on Twitter in the period between 2006 and 2017 showed that fake information spreads faster and wider than true information. They found that fake news is 70% more likely to be retweeted, therefore reaching people much faster. The effect of spreading fake news across different social media platform can be disastrous in many aspects of life. It can bias political campaigns and decisions, like what happened in the “Brexit” referendum [5], 2016 US elections, [6] and the recent 2020 US elections as well. One rumor can cause the stock market to lose millions of dollars. For example, a rumor about former US president Barack Obama being injured in an explosion cost the stock market millions of dollars [7]. With the emerge of the COVID-19 pandemic, lots of rumors spread over the social network about home remedies, off the shelf chemicals, and deadly side effects of new vaccines, which actually put lives at risk in believing those rumors [8].

In light of the problem presented, researchers have worked on the problem of detecting fake news on social media extensively over the past decade, especially with the advances in text classification techniques using deep learning and different natural language processing. The most prevalent approaches have focused on using the text of the social post to assess its veracity. In [9], a tree-structured recursive neural network was proposed to tackle the problem. A more complex approach based on CNN and RNN models was also presented in [10] for the identification of fake news on Twitter. In addition to traditional supervised learning approaches, unsupervised models were also used to attempt to approach the same problem [11]. A comprehensive survey for comparing the different machine learning approaches for tackling fake news detection was presented recently in [12].

With the emerge of the COVID-19 pandemic and the growing realization of the severe physical impact that misinformation on social media posed, creating health hazards and impeding the efforts to decrease the pandemic’s harm, more research has been focused on finding more efficient solutions to the problem. A deep learning model based on BERT [13] was presented in [14] that was trained on tweets about COVID-19. Another approach that incorporated sentiment analysis of tweets used to identify fake news about the pandemic was presented in [15].

The main limitation of all of the above models is the nature of the social media posts, which usually are short sentences, with some limitations on the total number of characters in some cases (e.g., Twitter). These short snippets of text do not usually contain enough information to tell if this news is real or not. Users can then easily be misled by these short, non-contextualized texts, thus believing and sharing them without proper validation if the post is fake or not. On the other hand, a news article holds more information than the title or the tweet posted about the news. However, users rarely search or read the full article searching for facts.

In this paper, we present a novel approach to tackle the problem of detecting fake news/posts on social media, coined as TChecker. The main hypothesis behind our proposed approach is that we can achieve higher performance in detecting fake news by incorporating the context of the social post as well as the content of the associated news article, thus enriching the model with more relevant signals for detecting the veracity of the post. Throughout the remainder of the paper, we will rely on Twitter as one of the most prevalent social platforms for sharing information. Our proposed approach, TChecker, uses BERTweet [16] for tweet representation and BERT for article representation, followed by a contextual layer of BiLSTM.

Our main research questions are as follows:

- RQ1: How would the introduction of contextual information enhance the veracity detection of social posts?
- RQ2: How would the incorporation of news articles alongside the post content enhance the veracity detection of social posts?

The paper is organized as follows. First, we present related work in the area of fake news detection focusing on social media, and then we explain our research methodology and the rationale behind our proposed model. Next, we present the experimental results of our proposed model on benchmark datasets and compare its performance with state-of-the-art models in fake news detection, and finally we conclude our work and present future directions.

2. Related Work

Detecting fake news is a problem tackled through different approaches that can be categorized mainly into a content-based approach and a social-based approach. In the content-based approach, the textual features are the main features, whereas in the social-based approach other features, including users' engagements, users' profile features, and network propagation features, are considered [17]. In this section, we are surveying the content-based approach and the social-based approach, focusing on the detection from social media.

2.1. Content-Based Approach

The most adopted approach in the task of identifying fake news is the content-based approach. In this approach, the textual features of the news are used in different models to identify the veracity of the news. This approach has been widely applied in detecting fake news from news posts and social posts.

Fake News Detection from News Articles

Verifying the truthfulness of news is a crucial step in the domain of publishing news, and checking the credibility of the source of information is an undeniable step in the publishing process. In the media domain, where journalists and others work in this domain, the verification of the news and its sources is their job. Journalists usually check the information against credible sources and verify this information is true before publishing it, that is, manual fact-checking.

With the increase in the volume of data roaming the internet every second, automatic techniques stepped in to help in the fact-checking process. Natural language processing and Information Retrieval techniques are applied to automatically identify fake news. A binary translating embedding (B-TransE) model was introduced in [18] to detect fake news based on a knowledge base graph; they evaluated their model by applying it to check the news in the dataset "Getting real about fake news" provided by Kaggle. CompareNet [19] is an end-to-end graph neural model that compares fake news against a knowledge base using entities.

The style-based approach relies on the content of the news post to detect its truthfulness based on the style of writing in the post. The style of writing would reveal the user's intention to post false or true information. The style of the writing is represented as features to be fed to the model for detecting the truthfulness of the post. This approach was used by [20–22] in deception detection, where deception is defined as the bad intention of authors to post intentionally false information. Those features were used in detecting fake news from news articles in [23,24].

Different machine learning algorithms have been applied to detect fake news from news posts through applying classification models on the textual content of the news. An analysis of different classifiers' performance on the LIAR dataset is presented in [25]; a comparison between the performance of Naïve Bayes, SVM, Random Forest, Logistic regression, and stochastic gradient classifier is presented, showing that the classifiers obtained near results except for the stochastic gradient classifier, which performed worse

than the others. Another comparison between the Naïve Bayes, Random Forest, passive aggressive, and LSTM models is presented in [26]; they applied the three models on a dataset consisting of 11,000 English articles labeled as fake or real. They showed that the passive aggressive classifier with TF-IDF representation could achieve the highest accuracy and F1 score, while the LSTM model could achieve almost the same accuracy, but it obtained a higher precision compared to that achieved by the passive aggressive classifier.

A classification model based on BiLSTM and self-attention layers is presented in [27] and applied on a dataset provided from Kaggle that consists of news articles labeled as fake or not. The articles are represented using GloVe [28] embeddings, then fed to the BiLSTM layer, followed by the self-attention layer, and then finally the classification layer. They compared their model to other models using different text representations, such as TF-IDF and BOW, and different neural networks, such as GRU, LSTM, and CNN.

Upon the introduction of BERT in 2019 as a pre-trained language model using deep bidirectional transformers, a major change in performance in NLP tasks occurred. BERT differs from other deep learning embeddings like word embedding and sentence/document embeddings, in that it takes into consideration the context of the word from the two directions of the sentence. This results in better representation of the words, capturing the contextual features of the text. Applying BERT in the task of detecting fake news has been introduced in many studies; in this section, we are going to present the significant results achieved by BERT-based models.

An evaluation of different language models, including BERT, RoBERTa [29], and DistilBERT [30], is presented in [31]; the authors also compared different architectures of neural network models, including a simple fully connected network, a CNN, and a combined CNN and RNN. They applied the models on different datasets of long and short text including Twitter datasets. Their study showed that simple neural network models can perform better than sophisticated models, and amongst the language models, RoBERTa performed slightly better on most of the datasets. However, all language models' performances were close to each other. Following the content-based approach, a BERT model is used followed by LSTM and fully connected layers, which is presented in [32]; they showed that vanilla BERT models could perform better than other content-based models. Moreover, adding the LSTM layer improved the performance of the model news titles of the Politifact dataset from those of the FakeNewsNet dataset [33]. Korean fake news was detected by [34] using a BERT-based model trained on the Korean language and BiLSTM for classification. A combination of three parallel blocks of 1D convolutional neural networks and a BERT model applied on news articles from a dataset collected during the 2016 US elections, provided by Kaggle, is presented in [35]. The model showed better performance than models using GloVe representations and other model architectures based on LSTM and CNN individually.

2.2. Fake News Detection from Social Media

Detecting fake news from social media is increasing these days. Different approaches are applied to detect fake news from social media posts in order to mitigate the harm caused by their spread. Early in 2016, a dataset was collected from Twitter by [36] at the time of five major events reported by journalists. A model of a convolutional neural network and LSTM was applied to this dataset and the results are presented in [10]. The LSTM model could achieve better results in terms of accuracy and F1 score than the CNN alone and the LSTM and CNN combined.

With the emergence of the COVID-19 pandemic, all researchers from different disciplines tried to find a way to reduce the effect of the pandemic worldwide. One way was to detect false information roaming social media. A dataset was collected and annotated from Twitter by [37] discussing different topics about the COVID-19 virus. They used Bag of Words (BOW) and n-grams to represent the text of tweets. They applied an ensemble of Naïve Bayes (NB), K Nearest Neighbor (KNN), Random Forest, function-sequential minimal optimization (SMO), and voted perceptron (VP). They found that the highest F1

score was achieved by the vote ensemble classifier. Different machine learning models which were applied to classify a set of collected tweets about COVID-19 are presented in [15]. The results showed that Random Forest could outperform other models like SVM, linear regression, and LSTM.

Different BERT-based models were used to detect misinformation from tweets specially regarding COVID-19. COVID-Twitter-BERT is introduced in [38]; it is a model of BERT pre-trained on a large corpus of tweets about COVID-19. The model is very topic-specific and can be useful in various NLP tasks related to representing tweets about COVID-19, such as detecting sentiments of tweets in [14]. Another COVID-19 dataset was collected from Twitter, and ensemble machine learning was applied in [39].

BERTweet followed by an output classification layer was used in [40] to detect misinformation from tweets. The model outperformed other text representation models such as GLoVe. BERTweet was compared with BERT cased and uncased pretrained models for detecting fake tweets about COVID-19 in [41]. BERTweet showed the best performance among different BERT models. For Arabic tweets, Ref. [42] presented a deep learning model based on ARaBERT, which is a BERT model trained on Modern Standard Arabic, to represent the tweets. The model uses the tweets' text and user features to detect the veracity of the tweets. BiLSTM and CNN networks were used for classification, where they showed close performance.

2.3. Social-Based Approach

In an attempt to understand more fake news through users' comments, an explainable fake news detection is presented in [17]. They use a model that relies on identifying comments that are explaining the core parts of the news article and how they are fake or not. The model is based on a co-attention network between the news articles and their users' comments. They also applied a ranking method to pick the comments with the most explanation. They compared their work on PolitiFact and GossipCop datasets to other content-based models using news articles only, and models that consider users' comments. They showed their model could achieve better results than the models in comparison.

Capturing features from comments and using them along with a two-level convolutional neural network learning representation from news content is presented in [43] as TCNN-URG. They use a conditional variational autoencoder for user comment generation to assist the news content classifier when user comments do not exist.

Integrating the features of the text of the news article, users' comment on it, and the source of the news is presented in [44] as a CSI model. Their model consists of three parts: the first one uses LSTM to capture the temporal representation of the article, the second part represents the user features, and the third part concatenates the results of the earlier parts into a classification model. Their experiments were performed on Twitter and Weibo datasets [45] and showed better results over content-based models.

TriFN is a tri-relationship embedding framework proposed by [46] that represents publisher–news relations and user–news interactions as embeddings and uses them together to detect fake news. They applied their model on PolitiFact and BuzzFeed datasets.

Incorporating the article's textual content, along with its creator and the subject of the news article, into a model presented as a deep diffusive network model is proposed in [47]. The latent features are extracted from the articles' text, their creators, and the subjects. The model uses a gated diffusive unit that accepts multiple inputs from different sources at the same time.

3. Methodology

In light of the reviewed literature, different approaches were presented to tackle the problem of detecting fake news from social media. The content-based approach relies mainly on the textual content of the tweets. However, due to the nature of Twitter, tweets usually are catchy posts without much detail. This nature makes it easier for users to go through lots of news and easily share it in no time. The spreading of misinformation mostly

comes from social media, as usually users do not fact-check posts before they share them. Trying to mimic what a fact-checker user would do, we retrieved the full news articles of the posted tweets, where more context can be useful in detecting the veracity of the tweet.

We started with a baseline model that follows the state of the art in content-based detection of fake tweets. The baseline starts with pre-processing the tweets by removing URLs and converting emojis into text, then uses BERTweet for content embedding, and then feeds the embedded vector into a deep neural network which consists of a fully connected layer and an output layer.

To answer the first research question, we investigated how adding BiLSTM to extract features from the BERTweet embeddings can enhance the results of veracity detection of tweets.

To answer the second research question, we enriched the tweets with their related news articles. Due to the completely different nature of the tweets' writing styles and limitations, and those of the professional news articles, our proposed model, TChecker, takes the tweets and their related news articles as separate inputs to two parallel networks. Each network consists of the embeddings and feature extraction phases. The model concatenates the embedded feature vector's results from the BiLSTM layers before passing them to the DNN part of the model. In the following subsections, we will illustrate the architecture of the described models.

3.1. Baseline Model

Fake news detection models have heavily relied on different techniques of representing news' textual content to detect its veracity, the most prominent of which is using a text embedding model (e.g., Google's BERT) to obtain a representative feature vector of the content.

With respect to detecting from social media, BERT has been used along with its variations, like RoBERTa, in [14,38,41,48]. BERTweet is a fine-tuned BERT-based model trained on English tweets. It showed enhancement in tweet classification tasks compared to other BERT models. It was used recently to detect fake news from Twitter, as shown in [40] and in an ensemble model in [49].

For our baseline, we followed the state of the art by using the tweet's text to create an embedding through BERTweet, followed by a fully connected layer reaching the output layer. The model architecture is shown in Figure 1.

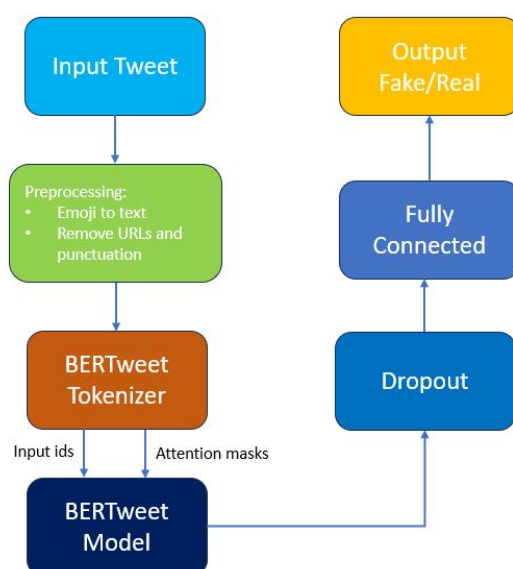


Figure 1. Baseline model architecture.

3.2. BERTweet BiLSTM Model

Rai et al. [32] recently showed that feeding the BERT embeddings to a deep neural network model composed of an LSTM and a number of fully connected layers showed better results than using BERT only on FakeNewsNet news titles. Following that approach, we added a BiLSTM layer after the BERTweet layer for better extraction of the features from the BERTweet embedded vectors. The architecture of the model is shown in Figure 2.

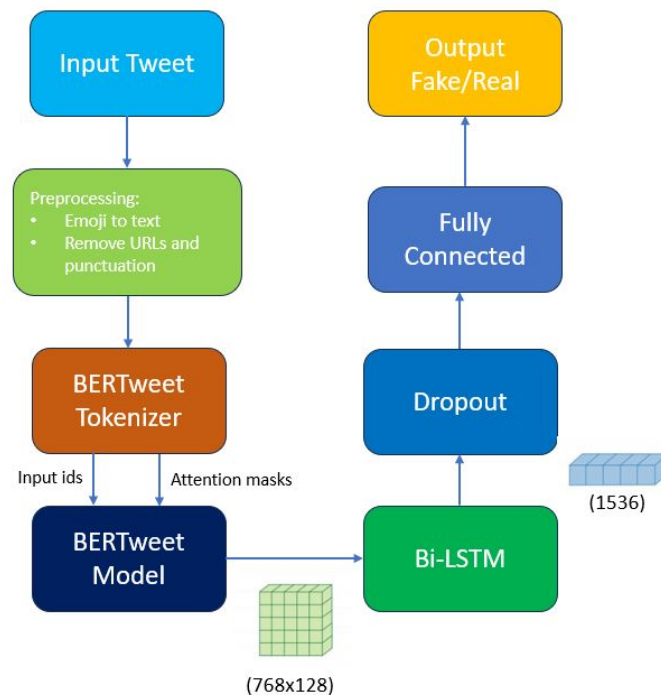


Figure 2. BERTweet–BiLSTM model architecture.

3.3. TChecker Model

We believe the full text of a news article will provide more signals for the machine learning model to be able to differentiate fake from authentic tweets. To explore the benefit of incorporating news articles in identifying fake tweets, we propose TChecker, a model that relies on two parallel BERT-based neural networks.

The model aims to capture the inherent differences in language, vocabulary, writing style, and limitations between news articles and users' tweets. Therefore, it generates separate feature embeddings for news articles and tweets through passing each input to a separate BERT/BERTweet model followed by a BiLSTM layer, then concatenating the two feature vectors and passing the concatenated vector to the fully connected layer. Separating the embeddings of the tweets from the actual news article helps preserve the unique characteristics of each, and concatenating both vectors as inputs to the neural network helps balance the importance of each source, without having intrinsic weighting based on the volume of the articles/tweets. The proposed model architecture is shown in Figure 3.

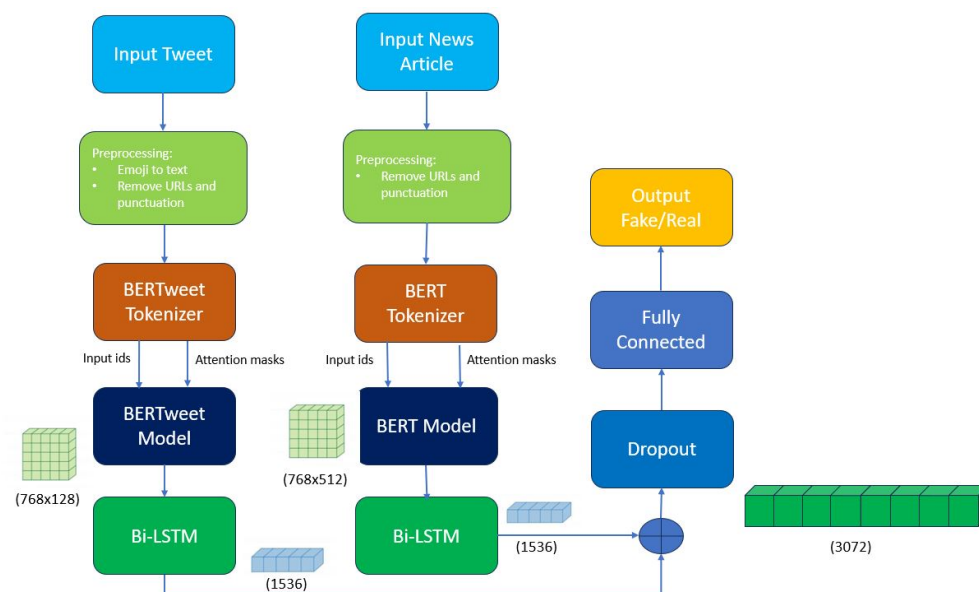


Figure 3. TChecker model architecture.

4. Experiments

In this section, we illustrate and compare the results of the baseline model and the proposed model on the described datasets. All experiments were conducted using PyTorch on Google Colab on an A100 GPU machine with 40 GB RAM. The code for the experiments can be found on github (<https://github.com/nadaaym/TChecker.git>, accessed on 10 November 2023).

Grid Search was used to find the best hyper-parameters including the number of epochs and the learning rate. The performance of the model showed stabilization after the second epoch. The exponential learning rate function was used such that the learning rate decreased every epoch, starting from 0.0001 till 0.0000006 with a factor of decay, $\gamma = 0.9$. Ten cross-fold validation was used to train and evaluate the model. The data was split into 80% training and 20% testing. Table 1 shows the model’s training parameters.

Table 1. Model training parameters.

Loss Function	Binary Cross Entropy
Optimizer	Adam
# Epochs	2
Batch Size	16
Learning Rate	0.0000006

4.1. Dataset

For our experiments, we used the FakeNewsNet datasets [33] that consists of two real life datasets, PolitiFact and GossipCop. The datasets were collected from fact-check websites and were manually revised and filtered. The related tweets to the news articles were collected using the Twitter API, along with their retweets, and user information. The FakeNewsNet dataset is considered a very rich dataset in terms of textual, social context, user information, network, and spatiotemporal information.

FakeNewsNet (PolitiFact)

There were 1056 news posts, which consist of titles, news content, and related tweets, with the amount of fake and real news. After investigating the dataset, we removed all entries that had missing news content and/or related tweets, which resulted in 666 entries.

FakeNewsNet (GossipCop)

There were 22,140 news posts, which consist of titles, news content, and related tweets, with the amount of fake and real news. After investigating the dataset, we removed all entries that had missing news content and/or related tweets, which resulted in 18,796 entries.

Since we are interested in detecting fake news from social media, we used the tweet-level dataset, where each entry is the tweet, the news article related to the tweet, and the label of the tweet as fake or real. There was a significant imbalance in the data between the number of fake and real tweets; thus, we balanced the datasets as shown in Table 2.

Table 2. Dataset statistics.

	Train			Test		
	Real	Fake	Total	Real	Fake	Total
PoliFact	60,664	60,664	121,328	11,370	11,370	22,740
GossipCop	248,286	248,286	496,572	81,056	81,056	162,112

4.2. Baseline Model

For the baseline model experiments, the tweets were preprocessed by removing urls and converting emojis into text, then represented as a vector using the BERTweet uncased model, resulting in 128 length vectors. This was followed by a fully connected layer of 768 neurons, then an output layer using the sigmoid function for classification. The results of the baseline model are shown in Table 3.

Table 3. Baseline results.

Data	Accuracy	Recall	Precision	F1-Score
Politifact	0.82	0.81	0.83	0.82
GossipCop	0.84	0.83	0.85	0.84

4.3. Effect of BiLSTM

In this experiment, we investigated the effect of adding BiLSTM to the model for better representation of the features from the embedded vectors. For this, a bidirectional LSTM layer of 768 units was added to the model to take the BERTweet embedded vector and output a feature vector of length 1536, which was fed to the fully connected layer of size 1536 reaching the output layer for classification. The results are shown in Table 4.

The results show how adding the bidirectional LSTM enhanced the performance of the model in detecting the veracity of tweets in both datasets.

Table 4. BERTweet–BiLSTM model results.

Model	Accuracy	Recall	Precision	F1-Score
Politifact Dataset				
Baseline	0.82	0.81	0.83	0.82
BERTweet–BiLSTM	0.85	0.85	0.85	0.85
GossipCop Dataset				
Baseline	0.84	0.83	0.85	0.84
BERTweet–BiLSTM	0.88	0.89	0.88	0.88

4.4. TChecker Model

For the proposed model, the news articles are embedded using a separate BERT–BiLSTM network with the same parameters. The feature vectors from the two networks are concatenated and fed to the rest of the network. Table 5 shows the results we were able to obtain for the baseline and the proposed model.

The results show that TChecker performed better than the baseline and the BERTweet–BiLSTM models. This emphasizes the role of incorporating the news article in detecting the veracity of tweets. The nature of news articles is different from the nature of tweets as they contain more information and linguistic features, which adds additional information that enhances the task of fake news detection from social media.

Table 5. TChecker model results.

Model	Accuracy	Recall	Precision	F1-Score
Politifact Dataset				
Baseline	0.82	0.81	0.83	0.82
BERTweet–BiLSTM	0.85	0.85	0.85	0.85
TChecker	0.93	0.93	0.93	0.93
GossipCop Dataset				
Baseline	0.84	0.83	0.85	0.84
BERTweet–BiLSTM	0.88	0.89	0.88	0.88
TChecker	0.91	0.91	0.91	0.91

In order to further validate our proposed model, we compared its performance to other content-based models as well other models in the literature that used both tweets and news articles, including the following:

- Text-CNN [50]: Text-CNN utilizes convolutional neural networks to model news articles, which can capture different granularities of text features with multiple convolution filters.
- TCNN-URG [43]: TCNN-URG consists of two major components: a two-level convolutional neural network to learn representations from news articles, and a conditional variational auto-encoder to capture features from user comments.
- CSI [44]: CSI is a hybrid deep learning model that utilizes information from the text, response, and source. The news representation is modeled via an LSTM neural network with the Doc2Vec embedding on the news articles and user comments as input.
- dEFEND [17]: This is a model uses the news articles and rank the user posts to select posts that explain why the post is real or fake.

We ran our experiments using the same two datasets for all the referenced models using their default parameters, and the results are summarized in Table 6. The results we obtained shows that TChecker outperforms all the other more complex models in the literature across all the evaluation criteria.

Table 6. TChecker performance compared to other models.

Input	Model	Accuracy	Recall	Precision	F1-Score
Politifact Dataset					
Articles	Text-CNN	0.653	0.863	0.678	0.76
Articles + Tweets	TCNN-URG	0.712	0.712	0.723	0.722
Articles + Tweets	CSI	0.807	0.813	0.821	0.817
Articles + Tweets	dEFEND	0.90	0.90	0.90	0.90
Articles + Tweets	TChecker	0.93	0.93	0.93	0.93
GossipCop Dataset					
Articles	Text-CNN	0.739	0.477	0.707	0.569
Articles + Tweets	TCNN-URG	0.736	0.521	0.715	0.603
Articles + Tweets	CSI	0.762	0.658	0.722	0.688
Articles + Tweets	dEFEND	0.808	0.808	0.808	0.808
Articles + Tweets	TChecker	0.91	0.91	0.91	0.91

5. Discussion

In the previous sections, we demonstrated the importance of incorporating contextual information well as complementing the social post's data with data from the associated news article in the detection of fake news/fact-checking social posts. Our initial hypothesis was that fake news identification should be looked at within the context of the tweets; thus, incorporating the contextual information in a traditional model like BERTweet should improve the performance. By incorporating a bi-directional LSTM model in addition to the use of BERTweet, we have shown that we were able to improve the performance of BERTweet on the benchmark FakeNewsNet dataset, thus validating our hypothesis. Additionally, given the nature and the limitations of the social posts (e.g., Twitter limiting tweet lengths to 140 characters), our second hypothesis is that we should be able to leverage the content of the news post itself to enhance the performance of our approach. This allows us to use the best of both worlds, rather than relying only on the news content, which in isolation might not be very indicative of whether the news is real or fake, or only on the social posts, which provide more signals but with limited content to analyze. Hence, we combined both sources along with the contextual signal in our proposed model, TChecker, and demonstrated that we could achieve a higher performance than the baseline models that use each source independently, as well as the models that just concatenate the inputs, through independent modeling of each source and combining the representative feature vectors (that capture the nuances of each source) for the final stage in our model. We believe one of the major contributions of our proposed approach is showing that we can build a generalized model that is able to capture the relevant signals independently from each source, and then combine the captured signals for teaching a joint model that is able to utilize these signals to come up with a final conclusion. The same approach, we believe, can be extended to incorporate even more sources if available, building the right pathway for extracting their relevant signals and enriching the results even further.

6. Conclusions

In this paper, we proposed TChecker, a deep learning model that enriches tweets' textual content with news articles for better detection of fake news on social media. We illustrated how incorporating the content of the news article related to a tweet can significantly help in the process of identifying its veracity. We also showed that the underlying differences in the nature of both sources necessitate a machine learning model that is able to address the dimensions and the nature of each source separately, as opposed to merely combining the text from both sources. Our experiments show that our proposed model is capable of modeling each data source in its correct dimensionality, balancing the importance of both the social post and the actual news content, thus achieving higher precision and recall in detecting fake tweets compared to a number of state-of-the-art models from the literature on both the PolitiFact and GossipCop datasets, which are part of the benchmark FakeNewsNet dataset.

The results obtained through the proposed model are promising and, more importantly, open the road for more ways to incorporate the wealth of information on social media, such as users' engagement, influence, demographics, and network propagation features, to develop more versatile and accurate models for assessing the veracity of posts on social media, which paves the road for more accurate and comprehensive automatic fact-checking tools to help users assess the quality of information they are exposed to on these platforms.

Author Contributions: Conceptualization, N.G. and H.S.; methodology, N.G. and H.S.; software, N.G.; validation, N.G.; formal analysis, N.G.; investigation, N.G.; resources, N.G.; data curation, N.G.; writing—original draft preparation, N.G.; writing—review and editing, N.G. and H.S.; visualization, N.G.; supervision, H.S. and A.R.; project administration, A.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding and the APC was funded by the American University in Cairo.

Data Availability Statement: Publicly available datasets were analyzed in this study. This data can be found here: <https://github.com/KaiDMML/FakeNewsNet>.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Boehm, L.E. The validity effect: A search for mediating variables. *Personal. Soc. Psychol. Bull.* **1994**, *20*, 285–293. [CrossRef]
2. Nickerson, R.S. Confirmation bias: A ubiquitous phenomenon in many guises. *Rev. Gen. Psychol.* **1998**, *2*, 175–220. [CrossRef]
3. Yuan, L.; Jiang, H.; Shen, H.; Shi, L.; Cheng, N. Sustainable Development of Information Dissemination: A Review of Current Fake News Detection Research and Practice. *Systems* **2023**, *11*, 458. [CrossRef]
4. Vosoughi, S.; Roy, D.; Aral, S. The spread of true and false news online. *Science* **2018**, *359*, 1146–1151. [CrossRef] [PubMed]
5. Pogue, D. How to Stamp Out Fake News. *Sci. Am.* **2017**, *316*, 24. [CrossRef] [PubMed]
6. Allcott, H.; Gentzkow, M. Social Media and Fake News in the 2016 Election. *J. Econ. Perspect.* **2017**, *31*, 211–236. [CrossRef]
7. Rapoza, K. Can ‘Fake News’ Impact the Stock Market? Section: Investing. Available online: <https://www.forbes.com/sites/kenrapoza/2017/02/26/can-fake-news-impact-the-stock-market/?sh=129496f92fac> (accessed on 10 September 2023).
8. Cinelli, M.; Quattrocchi, W.; Galeazzi, A.; Valensise, C.M.; Brugnoli, E.; Schmidt, A.L.; Zola, P.; Zollo, F.; Scala, A. The COVID-19 social media infodemic. *Sci. Rep.* **2020**, *10*, 16598. [CrossRef] [PubMed]
9. Ma, J.; Gao, W.; Wong, K.F. Rumor Detection on Twitter with Tree-structured Recursive Neural Networks. In Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Melbourne, Australia, 15–20 July 2018; Association for Computational Linguistics: Stroudsburg, PA, USA, 2018; pp. 1980–1989. [CrossRef]
10. Ajao, O.; Bhowmik, D.; Zargari, S. Fake News Identification on Twitter with Hybrid CNN and RNN Models. In Proceedings of the 9th International Conference on Social Media and Society, Melbourne, Australia, 15–20 July 2018; ACM: Copenhagen, Denmark, 2018; pp. 226–230. [CrossRef]
11. Yang, S.; Shu, K.; Wang, S.; Gu, R.; Wu, F.; Liu, H. Unsupervised Fake News Detection on Social Media: A Generative Approach. In Proceedings of the AAAI Conference on Artificial Intelligence, Honolulu, HI, USA, 27 January–1 February 2019; Volume 33, pp. 5644–5651. [CrossRef]
12. Hlaing, M.M.M.; Kham, N.S.M. Defining News Authenticity on Social Media Using Machine Learning Approach. In Proceedings of the 2020 IEEE Conference on Computer Applications (ICCA), Yangon, Myanmar, 27–28 February 2020; pp. 1–6. [CrossRef]
13. Devlin, J.; Chang, M.W.; Lee, K.; Toutanova, K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *arXiv* **2019**, arXiv:1810.04805.
14. Lin, H.Y.; Moh, T.S. Sentiment analysis on COVID tweets using COVID-Twitter-BERT with auxiliary sentence approach. In Proceedings of the 2021 ACM Southeast Conference, ACM SE ’21, Online, 15–17 April 2021; Association for Computing Machinery: New York, NY, USA, 2021; pp. 234–238. [CrossRef]
15. Jeyasudha, J.; Seth, P.; Usha, G.; Tanna, P. Fake Information Analysis and Detection on Pandemic in Twitter. *SN Comput. Sci.* **2022**, *3*, 456. [CrossRef]
16. Nguyen, D.Q.; Vu, T.; Tuan Nguyen, A. BERTweet: A pre-trained language model for English Tweets. In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations, Online, 16–20 November 2020; Association for Computational Linguistics: Stroudsburg, PA, USA, 2020; pp. 9–14. [CrossRef]
17. Shu, K.; Cui, L.; Wang, S.; Lee, D.; Liu, H. dDEFEND: Explainable Fake News Detection. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Anchorage, AK, USA, 4–8 August 2019; ACM: New York, NY, USA, 2019; pp. 395–405. [CrossRef]
18. Pan, J.Z.; Pavlova, S.; Li, C.; Li, N.; Li, Y.; Liu, J. Content Based Fake News Detection Using Knowledge Graphs. In Proceedings of the Semantic Web—ISWC 2018, Monterey, CA, USA, 8–12 October 2018; Vrandečić, D., Bontcheva, K., Suárez-Figueroa, M.C., Presutti, V., Celino, I., Sabou, M., Kaffee, L.A., Simperl, E., Eds.; Lecture Notes in Computer Science; Springer International Publishing: Cham, Switzerland, 2018; pp. 669–683. [CrossRef]
19. Hu, L.; Yang, T.; Zhang, L.; Zhong, W.; Tang, D.; Shi, C.; Duan, N.; Zhou, M. Compare to The Knowledge: Graph Neural Fake News Detection with External Knowledge. In Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers), Online, 1–6 August 2021; Association for Computational Linguistics: Stroudsburg, PA, USA, 2021; pp. 754–763. [CrossRef]
20. Siering, M.; Koch, J.A.; Deokar, A.V. Detecting Fraudulent Behavior on Crowdfunding Platforms: The Role of Linguistic and Content-Based Cues in Static and Dynamic Contexts. *J. Manag. Inf. Syst.* **2016**, *33*, 421–455. [CrossRef]
21. Zhang, D.; Zhou, L.; Kehoe, J.L.; Kilic, I.Y. What Online Reviewer Behaviors Really Matter? Effects of Verbal and Nonverbal Behaviors on Detection of Fake Online Reviews. *J. Manag. Inf. Syst.* **2016**, *33*, 456–481. [CrossRef]
22. Braud, C.; Søgaard, A. Is writing style predictive of scientific fraud? *arXiv* **2017**, arXiv:1707.04095.
23. Bond, G.D.; Holman, R.D.; Eggert, J.A.L.; Speller, L.F.; Garcia, O.N.; Mejia, S.C.; McInnes, K.W.; Ceniceros, E.C.; Rustige, R. ‘Lyn’ Ted’, ‘Crooked Hillary’, and ‘Deceptive Donald’: Language of Lies in the 2016 US Presidential Debates. *Appl. Cogn. Psychol.* **2017**, *31*, 668–677. [CrossRef]
24. Potthast, M.; Kiesel, J.; Reinartz, K.; Bevendorff, J.; Stein, B. A Stylometric Inquiry into Hyperpartisan and Fake News. *arXiv* **2017**, arXiv:1702.05638.

25. Agarwal, V.; Sultana, H.P.; Malhotra, S.; Sarkar, A. Analysis of Classifiers for Fake News Detection. *Procedia Comput. Sci.* **2019**, *165*, 377–383. [[CrossRef](#)]
26. Rohera, D.; Shethna, H.; Patel, K.; Thakker, U.; Tanwar, S.; Gupta, R.; Hong, W.C.; Sharma, R. A Taxonomy of Fake News Classification Techniques: Survey and Implementation Aspects. *IEEE Access* **2022**, *10*, 30367–30394. [[CrossRef](#)]
27. Mohapatra, A.; Thota, N.; Prakasam, P. Fake news detection and classification using hybrid BiLSTM and self-attention model. *Multimed. Tools Appl.* **2022**, *81*, 18503–18519. [[CrossRef](#)]
28. Pennington, J.; Socher, R.; Manning, C. Glove: Global Vectors for Word Representation. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), Doha, Qatar, 25–29 October 2014; Association for Computational Linguistics: Stroudsburg, PA, USA, 2014; pp. 1532–1543. [[CrossRef](#)]
29. Liu, Y.; Ott, M.; Goyal, N.; Du, J.; Joshi, M.; Chen, D.; Levy, O.; Lewis, M.; Zettlemoyer, L.; Stoyanov, V. RoBERTa: A Robustly Optimized BERT Pretraining Approach. *arXiv* **2019**, arXiv:1907.11692.
30. Sanh, V.; Debut, L.; Chaumond, J.; Wolf, T. DistilBERT, a distilled version of BERT: Smaller, faster, cheaper and lighter. *arXiv* **2020**, arXiv:1910.01108.
31. Anggrainingsih, R.; Hassan, G.M.; Datta, A. Evaluating BERT-Based Pre-Training Language Models for Detecting Misinformation. *arXiv* **2022**, arXiv:2203.07731.
32. Rai, N.; Kumar, D.; Kaushik, N.; Raj, C.; Ali, A. Fake News Classification using transformer based enhanced LSTM and BERT. *Int. J. Cogn. Comput. Eng.* **2022**, *3*, 98–105. [[CrossRef](#)]
33. Shu, K.; Mahudeswaran, D.; Wang, S.; Lee, D.; Liu, H. FakeNewsNet: A Data Repository with News Content, Social Context and Spatialtemporal Information for Studying Fake News on Social Media. *arXiv* **2019**, arXiv:1809.01286.
34. Lee, J.W.; Kim, J.H. Fake Sentence Detection Based on Transfer Learning: Applying to Korean COVID-19 Fake News. *Appl. Sci.* **2022**, *12*, 6402. [[CrossRef](#)]
35. Kaliyar, R.K.; Goswami, A.; Narang, P. FakeBERT: Fake news detection in social media with a BERT-based deep learning approach. *Multimed. Tools Appl.* **2021**, *80*, 11765–11788. [[CrossRef](#)] [[PubMed](#)]
36. Zubiaga, A.; Liakata, M.; Procter, R. Learning Reporting Dynamics during Breaking News for Rumor Detection in Social Media. *arXiv* **2016**, arXiv:1610.07363.
37. Olaleye, T.; Abayomi-Alli, A.; Adesemowo, K.; Arogundade, O.T.; Misra, S.; Kose, U. SCLAVOEM: Hyper parameter optimization approach to predictive modelling of COVID-19 infodemic tweets using smote and classifier vote ensemble. *Soft Comput.* **2022**, *27*, 3531–3550. [[CrossRef](#)]
38. Müller, M.; Salathé, M.; Kummervold, P.E. COVID-Twitter-BERT: A Natural Language Processing Model to Analyse COVID-19 Content on Twitter. *arXiv* **2020**, arXiv:2005.07503.
39. Dadgar, S.; Ghatee, M. Checkovid: A COVID-19 misinformation detection system on Twitter using network and content mining perspectives. *arXiv* **2021**, arXiv:2107.09768.
40. Kumar, A.; Jhunjunwala, N.; Agarwal, R.; Chatterjee, N. NARNIA at NLP4IF-2021: Identification of Misinformation in COVID-19 Tweets Using BERTweet. In Proceedings of the Fourth Workshop on NLP for Internet Freedom: Censorship, Disinformation, and Propaganda, Online, 6 June 2021; Association for Computational Linguistics: Stroudsburg, PA, USA, 2021; pp. 99–103. [[CrossRef](#)]
41. Kim, M.G.; Kim, M.; Kim, J.H.; Kim, K. Fine-Tuning BERT Models to Classify Misinformation on Garlic and COVID-19 on Twitter. *Int. J. Environ. Res. Public Health* **2022**, *19*, 5126. [[CrossRef](#)]
42. Alyoubi, S.; Kalkatawi, M.; Abukhodair, F. The Detection of Fake News in Arabic Tweets Using Deep Learning. *Appl. Sci.* **2023**, *13*, 8209. [[CrossRef](#)]
43. Qian, F.; Gong, C.; Sharma, K.; Liu, Y. Neural User Response Generator: Fake News Detection with Collective User Intelligence. In Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, Stockholm, Sweden, 13–19 July 2018; International Joint Conferences on Artificial Intelligence Organization: Stroudsburg, PA, USA, 2018; pp. 3834–3840. [[CrossRef](#)]
44. 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, Singapore, 6–10 November 2017; ACM: New York, NY, USA, 2017; pp. 797–806. [[CrossRef](#)]
45. Ma, J.; Gao, W.; Mitra, P.; Kwon, S.; Jansen, B.J.; Wong, K.F.; Cha, M. Detecting rumors from microblogs with recurrent neural networks. In Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence, IJCAI'16, New York, NY, USA, 9–15 July 2016; AAAI Press: New York, NY, USA, 2016; pp. 3818–3824.
46. Shu, K.; Wang, S.; Liu, H. Beyond News Contents: The Role of Social Context for Fake News Detection. In Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining, Melbourne, VIC, Australia, 11–15 February 2019; ACM: New York, NY, USA, 2019; pp. 312–320. [[CrossRef](#)]
47. Zhang, J.; Dong, B.; Yu, P.S. FakeDetector: Effective Fake News Detection with Deep Diffusive Neural Network. In Proceedings of the 2020 IEEE 36th International Conference on Data Engineering (ICDE), Dallas, TX, USA, 20–24 April 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 1826–1829. [[CrossRef](#)]
48. Alkhalifa, R.; Yoong, T.; Kochkina, E.; Zubiaga, A.; Liakata, M. QMUL-SDS at CheckThat! 2020: Determining COVID-19 Tweet Check-Worthiness Using an Enhanced CT-BERT with Numeric Expressions. *arXiv* **2020**, arXiv:2008.13160.

49. Kumar, A.; Singh, J.P.; Singh, A.K. COVID-19 Fake News Detection Using Ensemble-Based Deep Learning Model. *IT Prof.* **2022**, *24*, 32–37. [[CrossRef](#)]
50. Kim, Y. Convolutional Neural Networks for Sentence Classification. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), Doha, Qatar, 25–29 October 2014; Association for Computational Linguistics: Stroudsburg, PA, USA, 2014; pp. 1746–1751. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.