Mansoura Engineering Journal

Volume 48 | Issue 6 Article 9

October 2023

EHF-FNDM: An efficient hybrid features fake news detection methodology on social media

Haidy Samir Fahim
Mansoura,Dakahlia,Egypt, haidysamir251@gmail.com

Asmaa Mohamed Al-saied
dekerness-mansoura, Dakahlia,Egypt

Ahmed Shaban Samra
Electronics and Communications Engineering - Dep., Faculty of Engineering, Mansoura University.

Abeer Twakol Khalil
Electronics and Communications Engineering - Dep., Faculty of Engineering, Mansoura University.

Recommended Citation
Fahim, Haidy Samir; Al-saied, Asmaa Mohamed; Samra, Ahmed Shaban; and Khalil, Abeer Twakol (2023) "EHF-FNDM: An efficient hybrid features fake news detection methodology on social media," *Mansoura Engineering Journal*: Vol. 48 : Iss. 6 , Article 9.
Available at: https://doi.org/10.58491/2735-4202.3090

ORIGINAL STUDY

EHF-FNDM: An Efficient Hybrid Features Fake News Detection Methodology on Social Media

Haidy S. Fahim [a,*], Asmaa M. Al-Saied [b], Ahmed S. Samra [c], Abeer T. Khalil [c]

[a] Electronics and Communications Engineer, Mansoura University, Egypt
[b] Computer and Control System Department, Mansoura High Institute of Engineering and Technology, Egypt
[c] Electronics and Communications Engineering Department, Faculty of Engineering, Mansoura University, Egypt

Abstract

People are increasingly using social media to consume and share news. The inherent benefits of social media over traditional news media include its low cost and ease of access. In addition, publishing a news article requires less content censorship on social media. The rapid spread of 'fake news' on social media, that is, news that contains intentionally false information, has a significant negative impact on society. For instance, false information about the coronavirus disease '2019' has spread around the world like a virus. Therefore, developing effective methods to detect fake news early has great importance. In this paper, the (Efficient Hybrid Features Fake News Detection Methodology) EHF-FNDM model was proposed. It is a classification model for early detection of fake news based on hybrid features. This model was developed to identify fake news based on user profiles, tweets, and replies. It has a user model that can realize whether a user is spreading fake news or not. This user model was essential in determining whether or not the tweet was fake.

Keywords: Classification model, Deep learning, False information

1. Introduction

Social media has grown in popularity as a means of consuming news owing to the fact that it is free and easy to use, and posts can spread quickly. Therefore, it is an excellent way for people to obtain and publish various types of information (Shu et al., 2017, 2019). Because information spreads quickly on social media, detection mechanisms must be able to predict news quickly enough to prevent the spread of fake news. Therefore, detecting fake news on social media is a vital and technically difficult problem (Kaliyar et al., 2021). A survey paper on the issue of fake news detection presents a simple definition of fake news as a 'news article that is deliberately and verifiably untrue.' This definition is based on two main characteristics: intent and authenticity (Shu et al., 2017). Fake news has emerged as a major societal issue and a challenging task for social media companies to identify and stop. This has prompted some to take drastic measures, such as WhatsApp deleting two million of its users every month to stop the spread of false information (Vereshchaka et al., 2020). The most serious real-world impact is that false news appears to generate real-life fears. For example, a person going to a Washington, DC, pizzeria carrying an AR-15 rifle after reading online that 'this pizzeria was housing young children as sex slaves as part of a child abuse ring led by Hillary Clinton.' (https://www.nytimes.com/2016/12/05/business/media/comet-ping-pong-pizza-shooting-fake-news-consequences.html) Police later apprehended this man and charged him with firing an assault rifle in the restaurant (Kang and Goldman, 2016). Detecting and fighting fake news has involved human effort. Human experts manually evaluate the truthfulness of dialectical news stories on fact-checking websites such as Snopes (https://www.snopes.com/), Politifact

Received 4 June 2023; revised 24 August 2023; accepted 26 August 2023.
Available online 20 October 2023

* Corresponding author.
E-mail addresses: haidysamir251@gmail.com (H.S. Fahim), asma.m.alsaid@gmail.com (A.M. Al-Saied), shmed@mans.edu.eg (A.S. Samra), abeer.twakol@mans.edu.eg (A.T. Khalil).

https://doi.org/10.58491/2735-4202.3090

(https://www.politifact.com/), and Factcheck.org (https://www.factcheck.org/). The results of the judging are then made public as a reference for fact-checking. Manual fact-checking can assist readers in identifying fake news. But it falls short of the goal of early detection of fake news for the following reasons:

(1) Manual fact-checking takes time, which makes it difficult to identify and report fake news quickly.
(2) It cannot scale to deal with a significant volume of false information produced on the Internet (Liu and Wu, 2020).

With the rapid growth of machine learning and deep learning techniques in recent years (LeCun et al., 2015), automatic machine learning-based detection systems have emerged as a viable alternative to manual fact-checking (Liu and Wu, 2020). There is a key benefit to employing a deep learning model over traditional techniques. It does not need any handwritten features; alternatively, it determines the appropriate feature by itself (Kaliyar et al., 2021). The main contributions of this paper are summarized as follows:

(1) Twitter's API was used to obtain the missing values for the Gossipcop dataset from Twitter. This helps in creating a dataset with all the information needed to train the EHF-FNDM model.
(2) Developing a user model based on the users' profiles. The user model effectively helped in identifying fake news based on the user's previous tweets and user characteristics. It also helped the EHF-FNDM achieve better results.
(3) Concentrating on enhancing the efficiency of early fake news detection based on limited information. It has been demonstrated that the user who initially posted a tweet can be used to determine its veracity when it is just starting to spread. The text of the tweet, replies, retweets, and repliers' user profiles were also used to conduct more reliable predictions on the truthfulness of a tweet.
(4) Choosing the proper optimizer for the EHF-FNDM model. A comparison between the selected optimizer (Adam) and other optimizers was provided.
(5) Choosing the proper neural network for the EHF-FNDM model. The selected deep learning network (LSTM) was compared with other networks.

The following is how the paper is organized: Section II provides an overview of the relevant studies conducted on the research topic. The proposed method is covered in Section III. Section IV contains details on the results, the ablation study, and comparisons between different optimizers and deep neural networks. Section V contains a discussion. Section VI concludes the paper.

2. Literature survey

'Fake news' is a term used to describe news articles that are intentionally and verifiably false. There are three factors that led to the selection of this specific definition. First, the underlying intent of 'fake news' offers both theoretical and practical value, allowing for a deeper understanding and analysis of the topic. Second, any methods for confirming the veracity of information that are applicable to the limited definition of 'fake news' are also applicable to the broader definition. Third, this definition can clear up any confusion regarding fake news and relevant concepts (Shu et al., 2017). A brief summary of the work in the field of fake news detection is given in this section. Table 1 categorizes the features used by existing detection approaches.

Numerous studies already employ a straightforward method to identify fake news based on its content. The majority of current studies concentrate on the textual elements of news articles, such as the headlines of the news and body text, while only a few look at image or video (Jin et al., 2016). (Castillo et al., 2011) used a set of simple content-based features, such as emoticon symbols, question marks, words with a positive or negative attitude, and pronouns, to evaluate the veracity of information on Twitter. The main features of social context-based approaches are user-based, post-based, and network-based. User-based features are obtained from the profiles of user to measure the characteristics of users (Castillo et al., 2011). According to a recent study (Shu et al., 2018), certain individuals have different traits and are more inclined than others to propagate false information.

Castillo et al. (2011) employed a collection of core user-based characteristics to assess the veracity of the information supplied by its source user. Most of

Table 1. Categorization of features.

Feature category	Subcategory	Data source
News content-based	textual-based	headline, body text
	visual-based	video, image
Social context-based	user-based	user profile, user post history
	post-based	user comment, retweets
	network-based	diffusion network, social network

social media networks support these features, such as follower and friend count, as well as registration age. In order to obtain temporal-linguistic patterns from a series of user comments to recognize rumors, some studies (Chen et al., 2018) used deep learning techniques like recurrent neural networks (RNN). PTK, a graph kernel-based SVM classifier introduced by (Ma et al., 2017), detects high-order patterns distinguishing distinct forms of false news by comparing their propagation tree structures. Additionally, a lot of hybrid models exist that identify fake news using a variety of feature categories. A hybrid model that combines the text, response, and source features of a news article has been proposed by (Ruchansky et al., 2017).

3. The proposed EHF-FNDM methodology

In the proposed study, the EHF-FNDM model was implemented. The EHF-FNDM model contained another model that was developed to identify users who distribute fake news. This model was referred to as the user model. Eighteen user-based features were identified. These features were extracted from the user profile and post history to train the user model. Nine text-based features were extracted from the news content, and six features were extracted from the social context of a tweet. These features, as well as the credibility of the user feature which extracted from the user model, were used to train the EHF-FNDM model. Fig. 1 shows the architecture of the proposed work. The following is a summary of the steps involved:

(1) The popular fake news dataset Gossipcop and Twitter's API were used to obtain the missing values from Twitter.
(2) Then the dataset was prepared, and tokenization was used to convert the larger text into words.
(3) To train the user model, features from the source user profile were extracted. So, users' credibility can then be predicted.
(4) After that, the fake news model was trained. The user model and features extracted from tweet text, replies, retweets, and repliers were used.
(5) The suggested model's performance was evaluated using a variety of evaluation metrics.

3.1. Data collection

The open-source Gossipcop dataset (Shu et al., 2020), which is extracted from Twitter, was used. The main model is based on user characteristics, reply content, and user network. Twitter's advanced search API was used to collect the dataset's missing values from Twitter. Gossipcop contains a lot of data. The tweets that had no replies and whose user was suspended were deleted. The tweets that were no longer available were also deleted. Table 2 provides a description of the final dataset's statistics.

3.1.1. Data preprocessing

For preprocessing, the NLTK toolkit (https://www.nltk.org/), a commonly utilized open source NLP library, was used. It includes built-in methods and algorithms such as nltk.tokenize (in order to tokenize the text), nltk.stem.porter.Porter-Stemmer, and others. The following steps were taken to preprocess the dataset:

(1) Tokenization: It is the process of converting a stream of textual data into meaningful elements such as words, sentences, terms, symbols, or other tokens. Prior to the feature extraction process, it is the first step in natural language processing (Asghar et al., 2021). Tokenization can break up sentences, words, or sub-words. Two types of tokenization sentence tokenization and word tokenization were used. The process of dividing a text into words is known as word tokenization. For word tokenization, word tokenize from the nltk.tokenize package was used. Sentence tokenization is the process by which the text is divided into sentences. For sentence tokenization, PunktSentenceTokenizer and PunktTrainer from the nltk.tokenize.punkt module were used.
(2) Stemming: It is a technique for reducing words to their root (which is also referred to as lemma). The primary goal of stemming is to minimise the frequency of derived words. This goal was achieved using the Porter-Stemmer algorithm. Porter-Stemmer algorithms can be implemented to stem a word using the Porter-Stemmer class provided by NLTK. This class is capable of converting an input word into a final stem using a variety of standard word forms and suffixes. The resulting stem is frequently a shorter word with the same root meaning. Let's look at an example. For instance, terms like singing, sang, and singer will be reduced to its lemma, the word sing.

3.2. The proposed feature extraction layer

3.2.1. Social media terminologies

The following fundamental social media lingo that was used to describe the user characteristics was briefly explained:

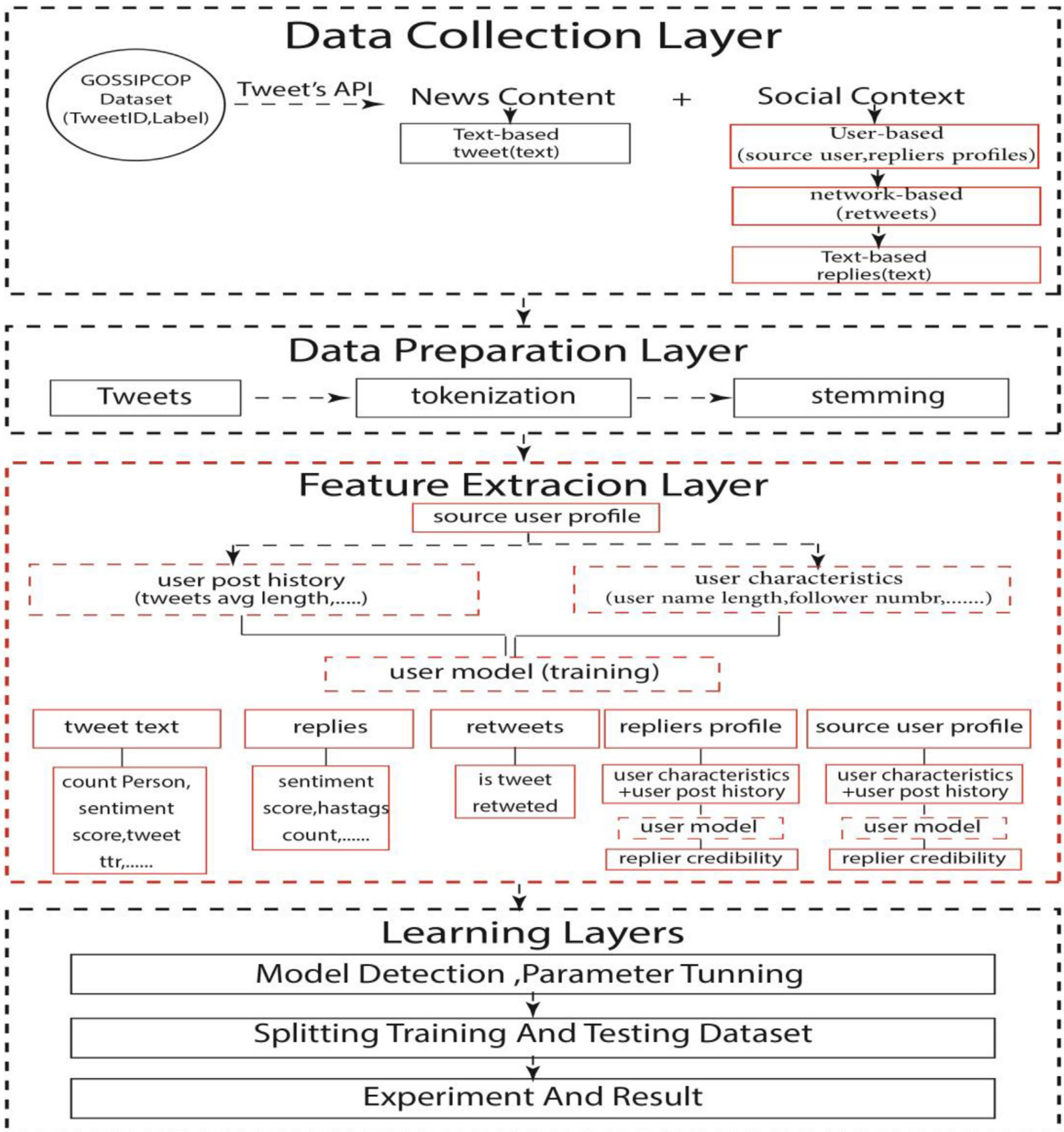


Fig. 1. The framework of the EHF-FNDM model.

(1) User: A human or computer program who creates an account on a social media platform.
(2) Follower: A user's follower is another user who has chosen to follow that user and who automatically receives all of that user's posts.

Table 2. Statistics of Gossipcop dataset.

Statistics of the Experimental	Gossipcop Dataset
News	617
Real news	329
Fake news	288

(3) Friend: A friend of a user is a different user who is being followed by the concerned user.
(4) Post: A post is any material published on social media by a user, such as a block of text, a picture, or a video.
(5) Retweeting: refers to the action of reposting or forwarded a message that has been published by another user.
(6) User characteristics: User characteristics are a group of features that characterize a user, such as the number of friends or statuses.

(7) Status: A social media post accompanied by information about the user who published it that was obtainable at the time the post was created.
(8) Source user: The user who was the first to share a news article on social media.
(9) Spreader: A news article's spreaders are the users who retweeted it.

3.2.2. The user model

To further enhance the EHF-FNDM detection approach, a deep learning model was created. This model takes into account users' historical behaviours and characteristics. Based on user characteristics and user post history, this model can determine whether a user is a fake news spreader or not. From the received user profile, user characteristics were first extracted. Then his or her previous tweets were collected. Because each user has a varied number of previous tweets, all the tweets had not been retrieved. Ten tweets were chosen, and if the user does not have any tweets, the algorithm can only make a judgement based on his user characteristics. Both situations were used to train the model. Few but effective features were chosen from the user's characteristics and past tweets. All the selected feature are shown in Table 3. These features include work of art average, screenname length, friend count, username length, follower count, and previous tweet lexical diversity, etc. A work of art is a NER feature and includes a book title, song title, etc. Lexical diversity can be a significant indicator for the detection of fake news, according to recent studies (Zhou and Zafarani, 2020). In Wikipedia, the lexical diversity is defined as: An aspect of 'lexical richness' is lexical diversity, which measures the ratio of various unique word stems (types) to the total number of words (tokens). This term is employed in applied linguistics, and it is quantitatively calculated via a variety of measurements, such as the Text-Type Ratio (TTR), vocd, and the measure of textual lexical diversity (MTLD).

It is a common issue with lexical diversity measures, particularly TTR, that text samples with many tokens produce lower TTR values because it is frequently necessary for the writer or speaker to reuse many function words. Newer lexical diversity measures make an effort to take text length sensitivity into account (https://en.wikipedia.org/wiki/Lexical_diversity). The Text-Type Ratio (TTR) was used as a measure of lexical diversity in the tweet text. The lex_div() function from the LexicalDiversity package was used for that purpose. To be more precise, this function counts the number of unique word types in the text of the news and divides that number by the total number of tokens. Keras (https://keras.io/), a Python wrapper of TensorFlow (https://www.tensorflow.org/), was used to implement the suggested model. The data splitting is shown in Table 4.

Adam was used to update the weights and bias. Adam, a technique for effective stochastic optimization that only needs first-order gradients and uses little memory. From estimates of the first and second moments of the gradients, the method calculates individual adaptive learning rates for various parameters; the name Adam is obtained from adaptive moment estimation. The method combines the benefits of two recently popular optimization techniques: RMSProp's (Root Mean Squared Propagation) ability to handle non-stationary objectives and AdaGrad's (Adaptive Gradient) ability to handle sparse gradients. The technique takes up little memory and is simple to use. In the field of machine learning, Adam was discovered to be reliable and well-suited to a variety of non-convex optimization problems (Kingma and Ba, 2014). To avoid overfitting, dropout was applied (Srivastava et al., 2014). Fig. 2 shows the structure of the model.

LSTM (Li and Wu, 2015) was used in the EHF-FNDM model. The detailed view of the proposed user model is shown in Fig. 3. The structure of the proposed user model started with a LSTM layer with 128 units and tanh as an activation function.

Table 3. The user characteristics obtained from Gossipcop user profiles.

No	Feature	Type
1	Username length	Integer
2	Screenname length	Integer
3	Personal description length	Integer
4	Followers count	Float
5	Friends count	Float
6	Listed count	Float
7	Has profile background tile	Boolean
8	Has profile background image	Boolean
9	Favorites count	Float
10	Statuses count	Float
11	Is account verified	Boolean
12	Has default profile	Boolean
13	Is GEO enabled	Boolean
14	Previous tweets lexical diversity ttr	Float
15	Previous tweets count words avg	Float
16	Previous tweets count word unique avg	Float
17	Previous tweets count sentence avg	Float
18	Previous tweets count work of art avg	Float

Table 4. Splitting of the experimental dataset.

Splitting of the dataset	
Training dataset	Testing dataset
499	118

```
Model: "sequential_1"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
lstm_2 (LSTM)                (None, 18, 128)           66560
_________________________________________________________________
dropout_1 (Dropout)          (None, 18, 128)           0
_________________________________________________________________
lstm_3 (LSTM)                (None, 64)                49408
_________________________________________________________________
dense_1 (Dense)              (None, 1)                 65
=================================================================
Total params: 116,033
Trainable params: 116,033
Non-trainable params: 0
_________________________________________________________________
```

Fig. 2. The structure of the user model.

Then a dropout layer was added. The dropout layer is a straightforward method for avoiding overfitting in neural networks. Dropout's main principle is to randomly remove units from the neural network during training to stop them from excessively co-adapting. Another LSTM layer was added with 64 neurons and a tanh activation function. The sigmoid function at the output layer predicts the class label that is, Fake class or Real class.

3.2.3. *The proposed EHF-FNDM model*

The user model has been trained, and it can now determine whether a user is spreading fake news or not. Therefore, the user model was used to extract two features (source user credibility and reply user credibility). Features from tweet text, including features based on Name Entity Recognition, tweet lexical diversity ttr, tweet sentiment score, were also retrieved. For the NER feature, the following details were taken from the tweet: PERSON, DATE, GPE (countries, cities, etc.), FAC (airports, buildings, etc.), PRODUCT, and ORDINAL (first, second etc.). The process of statistically determining whether a text is positive, negative, or neutral is known as sentiment analysis. 'Opinion mining' (https://www.geeksforgeeks.org/python-sentiment-analysis-using-vader/) is another name for it. To implement sentiment analysis, there are two main approaches. One is based on polarity, where texts are classified as either positive or negative, and the other is valence-based, where the strength of the sentiment is taken into account. The polarity_scores were calculated using

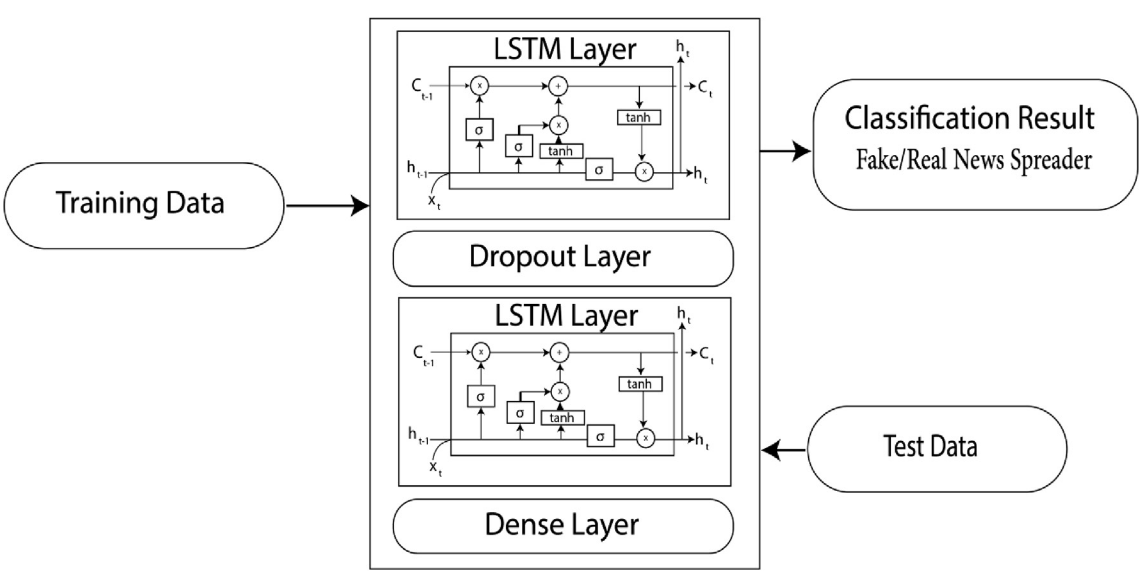


Fig. 3. Detailed view of the proposed user model.

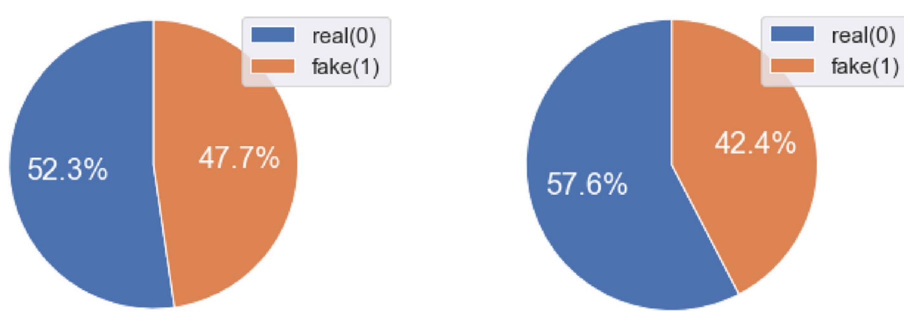
Table 5. List of features extracted from Gossipcop tweet and reply.

No	Feature	Type
1	Is tweet retweeted	Boolean
2	Tweet lexical diversity ttr	Float
3	Tweet sentiment score	Integer
4	Tweet count person	Integer
5	Tweet count date	Integer
6	Tweet count Gpe	Integer
7	Tweet count ORDINAL	Integer
8	Tweet count FAC	Integer
9	Tweet Count Product	Integer
10	Tweet Count Sentence	Integer
11	Tweet Source User Credibility	Boolean
12	Reply User Credibility	Boolean
13	Reply Text Event Count	Integer
14	Reply Text Person Count	Integer
15	Reply Hashtags Count	Integer
16	Reply Sentiment Score	Integer

the SentimentIntensityAnalyzer () function of the vaderSentiment package to apply the polarity-based method. Vader is generally used through the comments published on posts by social media users. However, any other kind of text can use this technique. As a result, it can be appropriate in the case of a tweet. When the rate was greater than 0.05, the text was classified as positive; when it was less than 0.05, it was classified as negative; and when it was in the middle, it was classified as neutral. The various degrees were defined as follows:

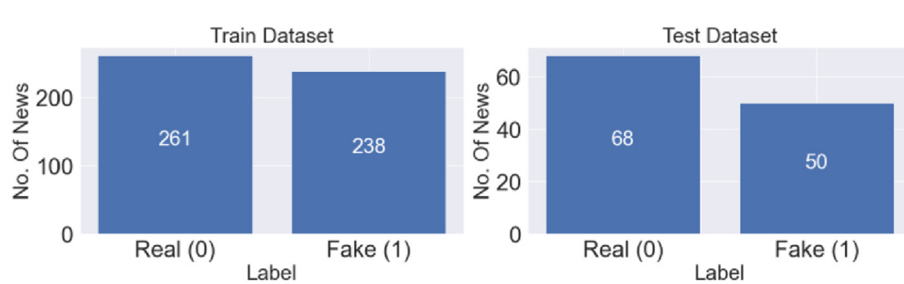
(1) Positive = 3
(2) Neutral = 2
(3) Negative = 1

Features like the number of reply hashtags, reply text events, and more that were extracted from the tweet's replies were used. The (Table 5) shows all the retrieved features. Data visualization was performed to present the data as frequency bar graphs and pie charts to gain insight into the dataset. Understanding the dataset's structure was the goal of the visualization. The pie chart for the training and testing datasets is shown in Fig. 4. Fig. 5 represents bar graphs for train and test dataset. The weights and bias were updated using Adam (Kingma and Ba, 2014). The LSTM neural network was used in



(a) Gossipcop train dataset (b) Gossipcop test dataset

Fig. 4. Pie chart of Gossipcop train and test dataset.



(a) Gossipcop train dataset (b) Gossipcop test dataset

Fig. 5. Bar plot of Gossipcop train and test dataset.

```
Model: "sequential"
_________________________________________________________________
Layer (type)                Output Shape              Param #
=================================================================
lstm (LSTM)                 (None, 16, 100)           40800
_________________________________________________________________
lstm_1 (LSTM)               (None, 16, 100)           80400
_________________________________________________________________
lstm_2 (LSTM)               (None, 50)                30200
_________________________________________________________________
dense (Dense)               (None, 1)                 51
=================================================================
Total params: 151,451
Trainable params: 151,451
Non-trainable params: 0
_________________________________________________________________
```

Fig. 6. The structure of the fake news model.

this model. The structure of the EHF-FNDM model is shown in Fig. 6.

4. The learning layer

Fig. 7 illustrates the structure of the EHF-FNDM model. The EHF-FNDM structure started with a two-layer LSTM with 100 neurons and Tanh as an activation function. Another LSTM layer with 50 units and Tanh as an activation function was added. Finally, the output layer was added with a sigmoid activation function that predicts the class label as fake or real.

5. Experiment results

Performance metrics such as accuracy, recall, precision and F1-score were taken into consideration for the evaluation. Here is a brief description of these metrics, where the number of True Positive results is represented by TP. False Positive results are indicated by FP, True Negative results are

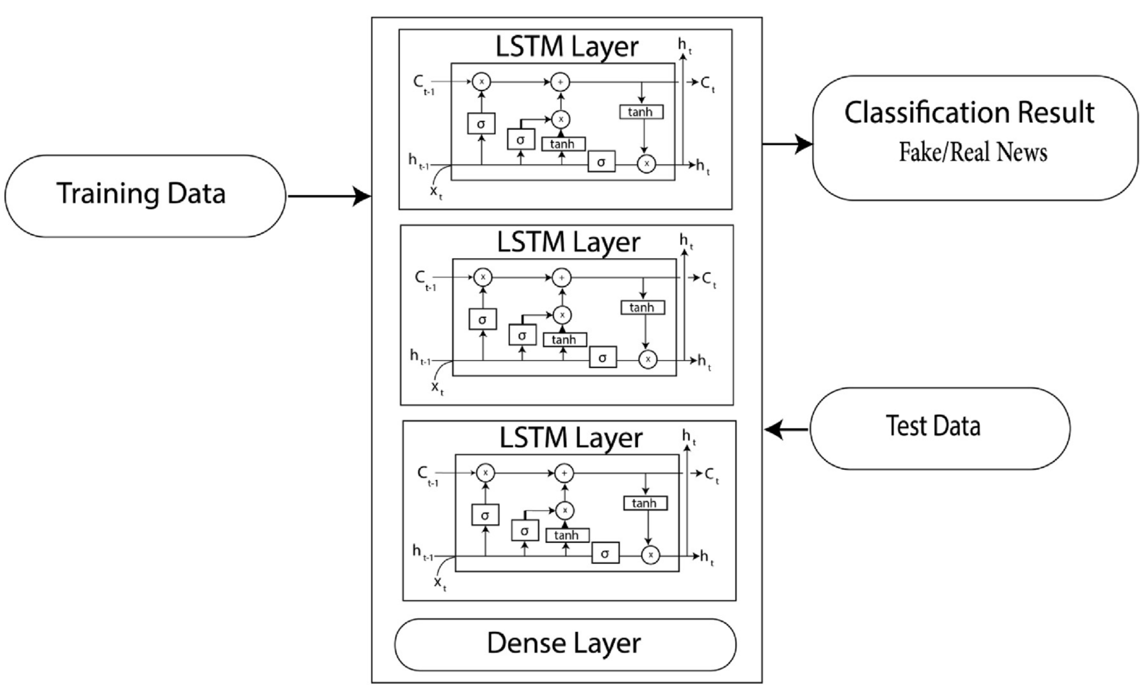


Fig. 7. Detailed view of the proposed EHF-FNDM model.

indicated by TN and FN is the number of False Negative results.

Accuracy is an estimate of the total number of instances that were correctly categorized and is calculated by:

$$\frac{TP + TN}{TP + FP + FN + TN} \tag{1}$$

Precision is measured as a percentage of relevant instances acquired from the overall number of instances using:

$$\frac{TP}{TP + FP} \tag{2}$$

Recall is computed as a percentage of relevant instances extracted from the overall number of relevant instances and calculated as follows:

$$\frac{TP}{TP + FN} \tag{3}$$

F1 Measure refers to the harmonic average of recall and precision provided by:

$$\frac{2 * precision * recall}{precision + recall} \tag{4}$$

5.1. The user model evaluation

The user model that was implemented to decide if the user spreads fake news or not achieved an accuracy of 0.8475, a precision of 0.86, a recall score of 0.85, and a F1-score of 0.85. The confusion matrix of the user model is shown in Fig. 8.

5.2. The EHF-FNDM model evaluation

The EHF-FNDM model was implemented to decide if the tweet is fake or real. This model used the tweet information, replies, and users' profiles. The EHF-FNDM model had an F1-score of 0.87, a recall score of 0.87, a precision score of 0.88, and an accuracy of 0.8729. Fig. 9 represents the model's confusion matrix. The time efficiency of the proposed framework for the test dataset (118 samples) is about 8.7 min The steps of news data, source user data, replier data, and reply data processing were done in a series. In the future, if the processing transforms from series to parallel, the processing time will be reduced. This means that the steps of news data, source user data, replier data, and reply data processing will start at the same time and be processed in parallel, which will decrease the processing time.

5.3. Ablation study

To better understand how each key component affects performance, a number of simplified versions of the suggested model were tested. One essential component was removed from each version. Here is a list of reduced internal models:

(1) EHF-FNDM - Usermodel: The user model was left out. Only news and replies were used to judge whether news was fake or not.
(2) EHF-FNDM - Usermodel-Replies: It relied solely on the news to determine whether it was fake or not.
(3) EHF-FNDM - Replies: No replies were included. To determine whether the news is fake or not, the news and the source's user information were used.

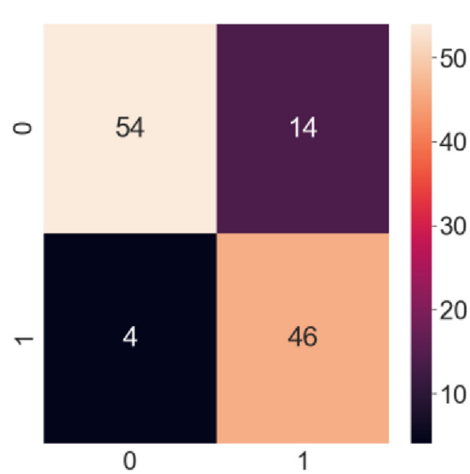


Fig. 8. The user model confusion matrix.

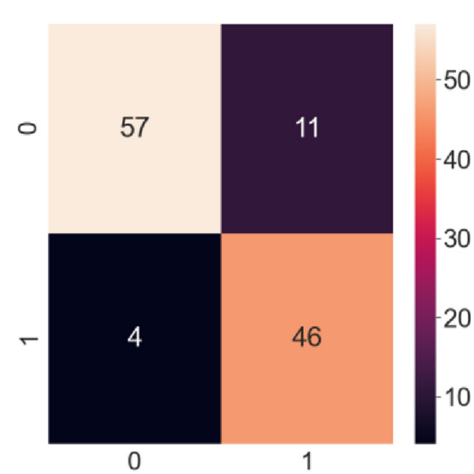


Fig. 9. Confusion matrix of the EHF-FNDM model.

Table 6. Performance comparison between full and reduced types of EHF-FNDM approach.

Approach	Accuracy	precision	Recall	F1-Score
EHF-FNDM - Usermodel	0.5847	0.62	0.58	0.58
EHF-FNDM - Usermodel-Replies	0.5508	0.61	0.55	0.54
EHF-FNDM - Replies	0.8559	0.87	0.86	0.86
EHF-FNDM	0.8729	0.88	0.87	0.87

(4) EHF-FNDM: The full model.

Table 6 compares the performance of the complete approach with the reduced types of the approach. The results showed that the EHF-FNDM performed worse when one important component was removed. The user model had the greatest of a significant impact on the detection accuracy, whereas the replies had the least.

5.4. Comparison of the chosen optimizer for the EHF-FNDM model (Adam) and other optimization algorithms

Finding the network parameters that minimize the loss function is an optimization process that is used to train a DNN. The Stochastic gradient descent (SGD) technique (Sutskever et al., 2013) is a basic deep learning algorithm that modifies the parameters iteratively based on the gradient for each training sample. SGD is less computationally complex than the original gradient descent method, which considers the entire dataset each time the parameters are changed. The hyperparameter learning rate controls the updating speed during the learning process. The concept of momentum is introduced in order to manage the oscillation of the SGD. This method, which is based on Newton's first law of motion, achieves quicker convergence and the right momentum, potentially improving SGD optimization results (Sutskever et al., 2013). Nesterov-Accelerated Gradient (Dozat, 2016) computes the gradient with momentum for the current position first, then leaps in the direction of the accumulated gradient and computes the value for the new position. Gradient is calculated for smaller steps that are either reduced to values below the current position or beyond the current position. This anticipatory updating stops SGD from convergenting too quickly, yielding better results.

The Adagrad optimization technique (Ward et al., 2020) adapts lower learning rates to frequently occurring features and higher learning rates to infrequent features. This technique's unique capability allows it to handle sparse data. The main advantage of using Adagrad is that the learning rate does not need to be changed manually; setting the initial learning rate to 0.01 causes the algorithm to adapt on its own. The disadvantage of Adagrad is the accumulation of squared gradients in the denominator, which results in an infinitesimally small learning rate. Adaptive Delta (Adadelta) (Zeiler, 2012) is an Adagrad extension that aims to strengthen Adagrad's weak points. Instead of accumulating all previous squared gradients, Adadelta limits the window of accumulating gradients to a fixed value, w. The sum of gradients is defined recursively as the decaying average of all previously squared gradients.

It is not necessary to set the learning rate with Adadelta. RMSprop (De et al., 2018) is an adaptive learning rate method that was presented around the same time as Adadelta. This was also created to address the radically diminishing problem of Adagrad. The adaptive learning rates for each parameter are calculated by adaptive moment estimation (ADAM) (Kingma and Ba, 2014).

Adam is simple to use, computationally effective, requires little memory, and is invariant to diagonal gradient rescaling. It is best suited for issues with a large quantity of data and/or parameters. It is also suitable for non-stationary goals and issues with extremely noisy and/or sparse gradients. The hyperparameters have straightforward interpretations and require little adjustment in most cases. Adam unites the features of two famous optimization methods: AdaGrad's ability to handle sparse gradients and RMSProp's ability to handle non-stationary targets. Adam stores the decaying average of previous squared gradients, similar to Adadelta and RMSprop, as well as the decaying average of past gradients, similar to Momentum. Overall, Adam was discovered to be robust and suitable for a wide variety of non-convex optimization issues in machine learning. Nesterov-Accelerated Gradient (NAG) outperforms Momentum. Adam and NAG are combined in Nesterov-Accelerated Adaptive Momentum Estimation (NADAM) (Kim and Kim, 2017).

A comparsion with other studies that used Adam optimizer is shown in Table 7. The SSLNews news classifier was proposed by (Konkobo et al., 2020). Using real-world datasets from Politifact and

Table 7. Comparison between the studies that used adam optimizer.

Reference	Accuracy	precision	Recall	F1-Score
Konkobo et al. (2020)	0.72	–	–	–
Vereshchaka et al. (2020)	0.75	–	–	–
Kumar et al. (2021)	0.72	–	–	0.71
Galli et al. (2022)	0.765	0.679	0.705	0.692
Truic et al. (2022)	0.53	–	–	0.32
Raza and Ding (2022)	0.748	0.724	0.776	0.749
EHF-FNDM	0.8729	0.88	0.87	0.87

Table 8. Comparison of the performances of optimization algorithms on Gossipcop dataset.

Optimizer	Accuracy	precision	Recall	F1-Score
SGD	0.5254	0.51	0.53	0.52
Adadelta	0.5593	0.55	0.56	0.55
Adagrad	0.6017	0.60	0.60	0.54
RMSprop	0.8559	0.86	0.86	0.86
Nadam	0.8644	0.87	0.86	0.87
Adam	0.8729	0.88	0.87	0.87

Gossipcop, they tested their model. SSLNews was able to attain accuracy ratings of 72.25% on Politifact and 70.35% on Gossipcop while using 25% of the data that had been labeled. When taking into account data obtained during the first 10 min of the news's dissemination, SSLNews had a Politifact accuracy of 71.10% and a Gossipcop accuracy of 68.07%. SSLNews is limited to using user-based features (Vereshchaka et al., 2020). Developed binary classifiers that use deep learning models such as long short-term memory and others to extract features from fake and true news. Utilising datasets that were extracted via the FakeNewsNet tool, it employed LSTM to obtain a 75% accuracy rate. It is limited to using content-based features. Moreover, it doesn't produce a higher result. In (Kumar et al., 2021), an XLNet fine-tuning model was suggested for predicting fake news in a multi-class and binary-class problem. It achieved an F1 of 0.71 and an accuracy of 0.72, both of which are lower than the EHF-FNDM result (Galli et al., 2022). Provides a framework for analysing and discussing the most extensively used machine and deep learning algorithms for detecting fake news. It also makes use of various feature combinations. This framework achieved an accuracy of 0.765, precision of 0.679, recall of 0.705, and F1 of 0.692. The result of this framework is less than the proposed model (Truic et al., 2022). suggested two Bidirectional Long Short-Term Memory (BiLSTM) structures with sentence transformers. The accuracy and F1-Score of a BiLSTM with BART (Bidirectional and Autoregressive Transformer) sentence transformer model on an English dataset are 0.53 and 0.32, respectively. Additionally, it can only use text and does not generate a better result. To identify false news (Raza and Ding, 2022), proposed the FND-NS model. It makes use of data from news items and social contexts. It is built using a Transformer architecture, which consists of two parts: an encoder to extract meaningful representations from fake news data and a decoder to forecast behaviour based on past observations. The FND-NS model has an accuracy of (74.8%), precision (72.4%), recall (77.6%), and F1 (74.9%). It does not obtain a higher result.

The performance of the various algorithms is summarized in Table 8. The obtained results demonstrated that the Adam algorithm presented the highest performance. The SGD algorithm presented the lowest performance.

5.5. Performance comparison of the chosen neural network (LSTM) model and other neural networks

The ability to process sequence data well distinguishes recurrent neural networks (RNN) from conventional artificial neural networks (ANN). Sequence data is the relationship between each word in a sentence and the words that come before and after it. As a result, RNN can be utilized to deal with natural language in context (Çöltekin and Rama, 2018). The gradient's role is to update the neural network's weight value. If the weight is too low, the gradient will disappear and the hidden layer close to the input layer won't be able to continue learning. If the weight is too large, the gradient will explode (Pascanu et al., 2013). As a result, this problem is handled by Long Short-Term Memory (LSTM) (Li and Wu, 2015), which provides memory blocks in its recurrent connections. Memory cells that store the network temporal states are present in each memory block. In order to control the flow of information, it also has gated units. Furthermore, the vanishing gradient problem can be significantly mitigated by residual connections in very deep networks (He et al., 2016). The GRU was presented by Cho et al. (2014). Table 9 provides a summary of the various deep learning networks' performance. According to the results, LSTM had the highest performance, whereas RNN had the lowest performance.

Table 9. Performance evaluation of deep learning networks using the Gossipcop dataset.

Deep learning networks	Accuracy	precision	Recall	F1-Score
SimpleRNN	0.8136	0.84	0.81	0.81
GRU	0.8390	0.84	0.84	0.84
Bidirectional (LSTM)	0.8390	0.84	0.84	0.84
LSTM	0.8729	0.88	0.87	0.87

6. Discussion

The method presented for this topic was not found in other research, so a comparison with the most recent research will be introduced. Fig. 10 shows an accuracy comparison between the different approaches and the proposed approach. AFARUL (Vereshchaka et al., 2020) created binary classifiers which extract features from fake and real news using deep learning models like long short-term memory (LSTM), recurrent neural network (RNN), and gated recurrent unit (GRU). AFARUL used datasets extracted by the FakeNewsNet (https://github.com/KaiDMML/FakeNewsNet) tool and achieved an accuracy of 75% through the use of LSTM. Two Bidirectional Long Short-Term Memory (BiLSTM) architectures with sentence transformers were proposed by FND_UBAST (Truic et al., 2022). On an English dataset, a BiLSTM with BART (Bidirectional and Autoregressive Transformer) sentence transformers model yields an accuracy of ~0.53. The news classifier SSLNews was suggested by (Konkobo et al., 2020). The SSLNews network is made up of three CNNs: one that is supervised, one that is unsupervised, and one that is shared. They evaluated their model using real-world datasets such as Politifact and Gossipcop. When using 25% of the data that had been labelled, SSLNews was able to achieve accuracy rates of 72.25% on Politifact and 70.35% on Gossipcop. SSLNews had a Politifact accuracy of 71.10% and a Gossipcop accuracy of 68.07% when considering data gathered during the first 10 min of the news's propagation. DANES (Truic et al., 2023) achieved an accuracy of 79.74% on BuzzFace via the model that uses a layer of LSTM, followed by a CNN layer for the social branch with MITTENS word embeddings and a layer of LSTM for the text branch.

FNAD_UR (Kumar et al., 2022) demonstrated a RoBERT-based model for detecting fake news articles. The experiment was done on a labeled test dataset provided by the organizer checkThat! of task 3, and an accuracy of 0.442 was observed. In NEW_FND (Oriola, 2021), features of news content are investigated to model n-gram, word embedding, and topic models as the base models and their hybrids. On the PolitiFact dataset, the n-gram model has an accuracy of 0.80. In (Abd Elaziz et al., 2023), a framework for detecting disinformation has been suggested in the context of the COVID-19 pandemic. This framework depends on multi-task learning (MTL) and meta-heuristic algorithms and achieved an accuracy of 59%. SV-FEND (Qi et al., 2023) is a new multi-modal detection model that uses cross–modal correlations to choose the most useful features and social context information for detection. SV-FEND achieved an accuracy of 79.31%. The proposed EHF-FNDM approach had an accuracy of 0.8729.

A comparison of the proposed approach's precision to other approaches is shown in (Fig. 11). On BuzzFace, DANES achieved a precision of 74.57%. NEW_FND had a precision of 0.79 on the PolitiFact dataset. The precision of HMLF was 53%.

FND_ELM (Truong and Tran, 2023) proposed an ensemble classification model based on using the collective wisdom of crowds to identify fake news. Without taking into account news content, fake news on Twitter is automatically detected by mining social interactions and the user's credibility. The method first extracts the features from a Twitter dataset, and then it uses a voting ensemble classifier made up of three classifiers—Support Vector Machine (SVM), Naive Bayes, and Softmax to divide news into fake and real news categories. It achieved a precision of 0.813. The precision of SV-FEND was 0.796. On the Gossipcop dataset, the proposed

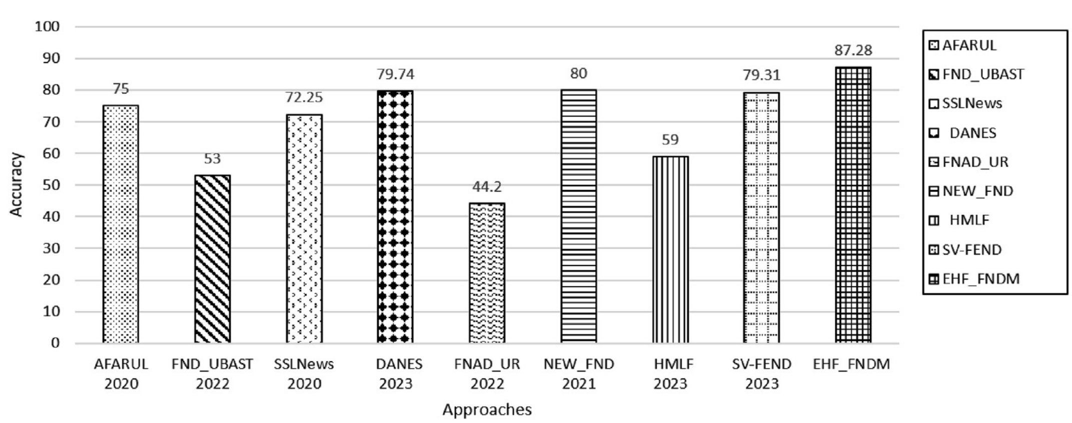


Fig. 10. Comparison between the proposed approach's accuracy and others.

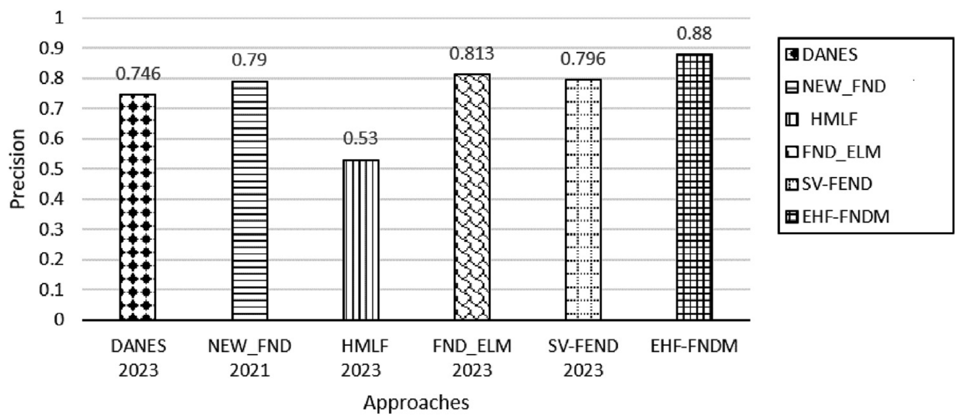


Fig. 11. Comparison between the proposed approach's precision and others.

approach EHF-FNDM achieved a precision of 0.88. Fig. 12 illustrates a recall comparison. DANES obtained a recall of 79.74% while NEW_FND had a recall of 0.78. HMLF achieved a 71% recall. A recall of 0.765 was attained by FND_ELM. SV-FEND achieved a recall of 0.793. The recall for the EHF-FNDM approach was 0.87.

Fig. 13 demonstrates a f1-score comparison. The f1-score for the FND_UBAST was 0.32, while the f1-score for the FNAD_UR was 0.296. A f1-score of 0.79 was attained by the NEW_FND. HMLF achieved a 53% F-measure. SV-FEND had a F1-score of 0.792. MFND (Suryavardan et al., 2023) presents the results of the Factify 2 shared task, which supplies a multi-modal fact verification and satire news dataset, as part of the DeFactify 2 workshop at AAAI'23. The data suggests a comparison-based approach to the task, with social media claims paired with supporting materials, both text and image, and split into five classes based on multi-modal relationships. In the second round of this task, the paper had over 60 participants and nine final test-set submissions. The best results were obtained when DeBERTa was used for text and Swinv2 and CLIP were used for images. The highest F1-score averaged across all five classes was 81.82%. FND_ELM achieved an F1-score of 0.788. The proposed approach had a f1-score of 0.87.

AFARUL does not achieve a better result. It can also only make use of content-based features. The diversity of fake news' content in terms of subject, presentation, and platform presents a common challenge for content-based detection techniques. Furthermore, news content features can be event-specific (Gupta et al., 2012; Sun et al., 2013). As a result, content-based features that perform well on one particular dataset of fake news may perform poorly on another (Liu and Wu, 2020). Furthermore, generalizability is a problem for machine learning models based on news content features (Tolosi et al., 2016). FND_UBAST is also confined to using only text and does not achieve a higher result. SSLNews can only employ user-based features. As a result,

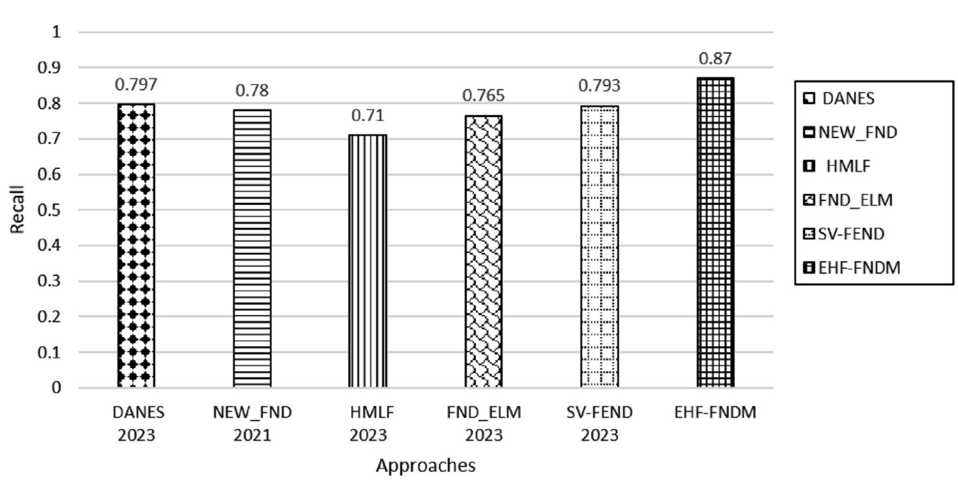


Fig. 12. Comparison between the proposed approach's recall and others.

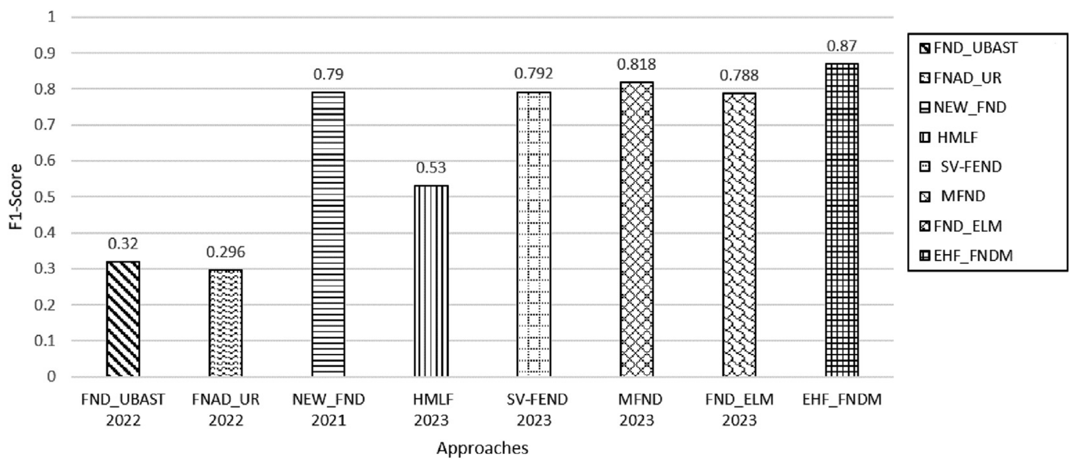


Fig. 13. Comparison between the proposed approach's f1-score and others.

user-based features alone cannot provide a complete picture of whether a news article is fake. However, user-based aspects of news spreaders, such as individuals who share or retweet a news story, may potentially provide us with more insight regarding the veracity of a news story (Liu and Wu, 2020). No higher result is attained with DANES. FNAD_UR is limited to using only text and does not achieve a higher result. New_FND can only use news content features. It also does not achieve a greater result. HMLF is restricted to using only contextual features and does not attain a higher result. FND_ELM and SV-FEND depend on the comments. One significant drawback of these methods is that, early in a news story's propagation process, there may not be many user comments, which can have a significant impact on model performance and easily lead to overfitting (Liu and Wu, 2020). They also did not achieve a higher result. MFND also did not attain a higher result.

7. Conclusion

In this study, the effectiveness of the proposed model for detecting fake news was demonstrated. Twitter's advanced search API was initially used to gather missing values for the dataset. Text-based features, user-based features, and post-based features were used in the hybrid model. A user model that can determine if a user is a fake news spreader or not was suggested. Results of the classification showed that the user model had a 0.8475 accuracy, a precision of 0.86, a 0.85 recall score, and an F1-score of 0.85. The accuracy, precision, recall, and F1-scores for the EHF-FNDM model were 0.8729, 0.88, 0.87, and 0.87, respectively. An ablation study was presented to test simplified versions of the model. The chosen optimizer was compared to others. The chosen deep learning network was also compared to other networks. The future direction of the study is to try to apply semi-supervised learning methods to overcome the issue of dataset availability. A paralization technique can also be applied to reduce the processing time.

Author credit statement

1. Conception or design of the work: Ahmed Shaban Samra, Abeer Twakol Khalil, Asmaa Mohamed Al-saied,Haidy Samir Fahim (25%/25%/25%/25%). 2. Data collection and tools: Ahmed Shaban Samra, Abeer Twakol Khalil, Asmaa Mohamed Al-saied,Haidy Samir Fahim (25%/20%/30%/25%). 3. Data analysis and interpretation: Ahmed Shaban Samra, Abeer Twakol Khalil, Asmaa Mohamed Al-saied,Haidy Samir Fahim (20%/30%/20%/30%). 4. Methodology: Ahmed Shaban Samra, Abeer Twakol Khalil, Asmaa Mohamed Al-saied,-Haidy Samir Fahim (30%/20%/25%/25%). 5. Project administration: Ahmed Shaban Samra, Abeer Twakol Khalil, Asmaa Mohamed Al-saied,Haidy Samir Fahim (25%/25%/25%/25%). 6. Software:Ahmed Shaban Samra, Abeer Twakol Khalil, Asmaa Mohamed Al-saied,Haidy Samir Fahim (25%/25%/25%/25%). 7. Drafting the article: Ahmed Shaban Samra, Abeer Twakol Khalil, Asmaa Mohamed Al-saied,Haidy Samir Fahim (20%/30%/25%/25%). 8. Critical revision of the article: Ahmed Shaban Samra, Abeer Twakol Khalil, Asmaa Mohamed Al-saied,Haidy Samir Fahim (30%/25%/25%/20%).

Authored by: Haidy Samir Fahim Ibrahim. Master student, Dept. of Electronics and Communications Engineering, Faculty of Engineering, Mansoura University. **Supervised by: Prof. Dr. Ahmed Shaban Samra**. Professor, Dept. of Electronics and

Communications Engineering, Faculty of Engineering, Mansoura University. **Assoc. Prof. Dr.Abeer Twakol khalil**. Associate professor, Dept. of Electronics and Communications Engineering, Faculty of Engineering, Mansoura University. **Dr. Asmaa Mohamed Al-saied**. Assistant Prof., computer and control system department, works at Mansoura High Institute of Engineering and technology.

Funding statement

There is no funding for the researcher.

Conflicts of interest

There is no conflict of interest.

References

Abd Elaziz, M., Dahou, A., Orabi, D.A., Alshathri, S., Soliman, E.M., Ewees, A.A., 2023. A hybrid multitask learning framework with a fire hawk optimizer for Arabic fake news detection. Mathematics 11, 258.

Asghar, M.Z., Subhan, F., Ahmad, H., Khan, W.Z., Hakak, S., Gadekallu, T.R., Alazab, M., 2021. Senti-eSystem: a sentiment-based eSystem-using hybridized fuzzy and deep neural network for measuring customer satisfaction. Software Pract. Ex. 51 (3), 571–594.

Castillo, C., Mendoza, M., Poblete, B., 2011. Information credibility on twitter. In: Proceedings of the 20th International Conference on World Wide Web, pp. 675–684.

Chen, T., Li, X., Yin, H., Zhang, J., 2018. Call attention to rumors: deep attention based recurrent neural networks for early rumor detection. In: Trends and Applications in Knowledge Discovery and Data Mining: PAKDD 2018 Workshops, BDASC, BDM, ML4Cyber, PAISI, DaMEMO, Melbourne, VIC, Australia, June 3, 2018, Revised Selected Papers 22. Springer, pp. 40–52.

Cho, K., Van Merriënboer, B., Bahdanau, D., Bengio, Y., 2014. On the properties of neural machine translation: encoder-decoder approaches. arXiv preprint arXiv 1409, 1259.

Çöltekin, Ç., Rama, T., 2018. Tübingen-oslo at SemEval-2018 task 2: SVMs perform better than RNNs in emoji prediction. In: Proceedings of the 12th International Workshop on Semantic Evaluation, pp. 34–38.

De, S., Mukherjee, A., Ullah, E., 2018. Convergence guarantees for RMSProp and ADAM in non-convex optimization and an empirical comparison to Nesterov acceleration. arXiv preprint arXiv 1807, 06766.

Dozat, T., 2016. Incorporating Nesterov Momentum into Adam.

Galli, A., Masciari, E., Moscato, V., Sperlí, G., 2022. A comprehensive Benchmark for fake news detection. J. Intell. Inf. Syst. 59 (1), 237–261.

Gupta, M., Zhao, P., Han, J., 2012. Evaluating Event Credibility on Twitter. Proceedings of the 2012 SIAM International Conference on Data Mining. SIAM, pp. 153–164.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778. https://doi.org/10.1109/CVPR.2016.90.

Jin, Z., Cao, J., Zhang, Y., Zhou, J., Tian, Q., 2016. Novel visual and statistical image features for microblogs news verification. IEEE Trans. Multimed. 19 (3), 598–608. https://doi.org/10.1109/tmm.2016.2617078.

Kaliyar, R.K., Goswami, A., FakeBERT, Narang P., 2021. Fake news detection in social media with a BERT-based deep learning approach. Multimed. Tool. Appl. 80 (8), 11765–11788 . https://doi.org/10.1007/s11042-020-10183-2.

Kang, C., Goldman, A., 2016. In Washington pizzeria attack, fake news brought real guns. N. Y. Times 5, A1.

Kim, J., Kim, H., 2017. An effective intrusion detection classifier using long short-term memory with gradient descent optimization. In: 2017 International Conference on Platform Technology and Service (PlatCon). IEEE, pp. 1–6.

Kingma, D.P., Ba, J., 2014. Adam: a method for stochastic optimization. arXiv preprint arXiv 1412, 6980.

Konkobo, P.M., Zhang, R., Huang, S., Minoungou, T.T., Ouedraogo, J.A., Li, L., 2020. A deep learning model for early detection of fake news on social media. In: 2020 7th International Conference on Behavioural and Social Computing (BESC). IEEE, pp. 1–6.

Kumar, A., Trueman, T.E., Cambria, E., 2021. Fake news detection using XLNet fine-tuning model. In: 2021 International Conference on Computational Intelligence and Computing Applications (ICCICA). IEEE, pp. 1–4.

Kumar, S., Kumar, G., Singh, S.R., 2022. TextMinor at CheckThat! 2022: Fake News Article Detection Using RoBERT. In: Working Notes of CLEF 2022 – Conference and Labs of the Evaluation Forum. CLEF, Bologna, Italy.

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521 (7553), 436–444. https://doi.org/10.1038/nature14539.

Li, X., Wu, X., 2015. Constructing long short-term memory based deep recurrent neural networks for large vocabulary speech recognition. In: 2015 Ieee International Conference on Acoustics, Speech and Signal Processing (Icassp). IEEE, pp. 4520–4524.

Liu, Y., Wu, Y.-F.B., 2020. Fned: a deep network for fake news early detection on social media. ACM Trans. Inf. Syst. 38, 1–33.

Ma, J., Gao, W., Wong, K.F., 2017. Detect Rumors in Microblog Posts Using Propagation Structure via Kernel Learning. Association for Computational Linguistics.

Oriola, O., 2021. Exploring N-gram, word embedding and topic models for content-based fake news detection in FakeNewsNet evaluation. Int. J. Comput. Appl. 975, 8887.

Pascanu, R., Mikolov, T., Bengio, Y., 2013. On the difficulty of training recurrent neural networks. In: International Conference on Machine Learning, vol. 28. Pmlr, Atlanta, Georgia, USA, pp. 1310–1318.

Qi, P., Bu, Y., Cao, J., Ji, W., Shui, R., Xiao, J., et al., 2023. FakeSV: a multimodal benchmark with rich social context for fake news detection on short video platforms. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 37, pp. 14444–14452. https://doi.org/10.1609/aaai.v37i12.26689.

Raza, S., Ding, C., 2022. Fake news detection based on news content and social contexts: a transformer-based approach. Int J Data Sci Anal 13 (4), 335–362. https://doi.org/10.1007/s410.

Ruchansky, N., Seo, S., Liu, Y., 2017. Csi: a hybrid deep model for fake news detection. In: Proceedings of the 2017 ACM on Conference on Information and Knowledge Management, pp. 797–806.

Shu, K., Sliva, A., Wang, S., Tang, J., Liu, H., 2017. Fake news detection on social media: a data mining perspective. ACM SIGKDD Explor Newsl 19, 22–36.

Shu, K., Wang, S., Liu, H., 2018. Understanding user profiles on social media for fake news detection. In: 2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR). IEEE, pp. 430–435.

Shu, K., Wang, S., Liu, H., 2019. 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 (WSDM '19), pp. 312−320.
Shu, K., Mahudeswaran, D., Wang, S., Lee, D., Liu, H., 2020. Fakenewsnet: a data repository with news content, social context, and spatiotemporal information for studying fake news on social media. Big Data 8, 171−188.
Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R., 2014. Dropout: a simple way to prevent neural networks from overfitting. J Mach LearnResh 15 (1), 1929−1958.
Sun, S., Liu, H., He, J., Du, X., 2013. Detecting Event Rumors on Sina Weibo Automatically. Web Technologies and Applications: 15th Asia-Pacific Web Conference, APWeb 2013. April 4-6, 2013. Proceedings 15. Springer, Sydney, Australia, pp. 120−131.
Suryavardan, S., Mishra, S., Chakraborty, M., Patwa, P., Rani, A., Chadha, A., et al., 2023. Findings of factify 2: multimodal fake news detection. arXiv preprint arXiv 2307, 10475.
Sutskever, I., Martens, J., Dahl, G., Hinton, G., 2013. On the importance of initialization and momentum in deep learning. In: International Conference on Machine Learning, vol. 28. PMLR, Atlanta, Georgia, USA, pp. 1139−1147.
Tolosi, L., Tagarev, A., Georgiev, G., 2016. An analysis of event-agnostic features for rumour classification in twitter. In: Proceedings of the International AAAI Conference on Web and Social Media, pp. 151−158.
Truică, C.-O., Apostol, E.-S., Paschke, A., 2022. Awakened at CheckThat! 2022: fake news detection using BiLSTM and sentence transformer. Working Notes of CLEF.
Truică, C.-O., Apostol, E.-S., Karras, P., 2023. DANES: deep neural network ensemble architecture for social and textual context-aware fake news detection. arXiv preprint arXiv 2302. arXiv:2302.01756v1 [cs.AI].
Truong, H.B., Tran, V.C., 2023. A framework for fake news detection based on the wisdom of crowds and the ensemble learning model. Comput. Sci. Inf. Syst. 48−48. https://doi.org/10.2298/CSIS230315048T.
Vereshchaka, A., Cosimini, S., Dong, W., 2020. Analyzing and distinguishing fake and real news to mitigate the problem of disinformation. Comput. Math. Organ. Theor. 26, 350−364.
Ward, R., Wu, X., Bottou, L., 2020. Adagrad stepsizes: sharp convergence over nonconvex landscapes. J. Mach. Learn. Res. 21, 9047−9076.
Zeiler, M.D., 2012. Adadelta: an adaptive learning rate method. CoRR, abs/1212.5701. Available at: http://arxiv.org/abs/1212.5701.
Zhou, X., Zafarani, R., 2020. A survey of fake news: fundamental theories, detection methods, and opportunities. ACM Comput. Surv. 53 (5), 1−40. Available at: http://arxiv.org/abs/1212.5701.