Research Highlights

- This paper proposes a Hashtag Context- aware Fake News Detection **HCFND** for fake news detection.
- **HCFND** extract information from posts under hashtags as external context for fake news detection.
- **HCFND** outperforms existing state-of-the-art models from literature in fake news detection over social media.

Pattern Recognition Letters
journal homepage: www.elsevier.com

Fake News Detection using Hashtag context

Sujit Kumar[a,**], Shifali Agrahari[a], Priyank Soni[a], Aayush Sachdeva[a], Sanasam Ranbir Singh[a]

[a]*Indian Institute of Technology Guwahati, Guwahati, 781039, Assam, India*

ABSTRACT

The proliferation of social media platforms has resulted in an exponential increase in user-generated content, facilitating the rapid and widespread dissemination of information. However, this ease of content sharing has also enabled the propagation of false or misleading information, often called "fake news," which can have detrimental consequences for society. Existing studies in the literature rely on content in source posts, social interaction networks, and external evidence to verify the authenticity of the posts. However, studies in the literature fail to detect the following case. (i) Sparsity and limited words in social media posts heavily affect the performance of content-based methods. (ii) Social interaction-based methods require a huge social interaction network for a given source post, which is easily only available for every social media post. (iii) Social media discussions sometimes precede or surpass mainstream media reporting and information from external sources such as Knowledge Base and Wikipedia. Consequently, in such circumstances, it isn't easy to get external information that will help verify the authenticity of social media posts. To address the above-mentioned limitations, this, study proposes *Hashtag Context-aware Fake News Detection* (HCFND). Our proposed model, HCFND, leverages information posted under the same hashtag and relevant posts extracted from named entities mentioned in the source post as external sources of information to enrich the data from the same community. We conduct our experiment over two publicly available benchmark datasets. Our experimental results suggest that our proposed model outperforms state-of-the-art methods in the literature.

1. Introduction

Social media, with its vast reach and accessibility, has become a significant platform for the rapid spread of fake news, impacting critical areas like healthcare and politics. The pervasive influence of misinformation erodes trust in institutions, fostering skepticism and polarization within societyAli et al. (2022) Zhou & Zafarani (2020). Beyond societal implications, fake news can manipulate public opinion, disrupt democratic processes, and influence economic factors during crises. Recognizing these far-reaching consequences, the detection of fake news on social media becomes imperative to counter the spread of misinformation and uphold the integrity of information channels. Earlier studies on fake news detection can be grouped into three distinct groups, namely: *Content-based*, *social interaction network* and *evidence claim verification*. Content-based fake news detection aims to identify false information by analyzing the content of social media posts. Content-based methods either employ bag-of-words features with traditional machine learning techniques Choudhary & Arora (2021); Hakak et al. (2021); Castillo et al. (2011); Popat (2017) or sequential encoding method Ma et al. (2016); Ahn & Jeong (2019); Potthast et al. (2018); Karimi et al. (2018); Bal et al. (2020); Hamed et al. (2023); Yu et al. (2017); Jain et al. (2023) over social media posts. As reported in the study Yan et al. (2015), sparsity and limited words of present in social media posts heavily affect the performance of content-based methods. Social interaction network-based methods evaluate the social connections, user profiles, community structure and temporal patterns in the news spreading. Social interaction network-based methods involve understanding social connections, user profiles Lu & Li (2020); Nguyen et al. (2020); Zhou et al. (2022), community structure Ruchansky et al. (2017); Wang et al. (2022); Shahid et al. (2022) and temporal patterns in the tweet propagation network Shu et al. (2020) or correlation between user

**Corresponding author:
e-mail: sujitkumar@iitg.ac.in (Sujit Kumar)

profiles and user interaction Lu & Li (2020); Nguyen et al. (2020); Zhou et al. (2022); Cheng et al. (2021) within social interaction network. Social interaction network based requires hug social interaction network to determine the authenticity of social media post, which is not available for every social media post Yang et al. (2022). Evidence claim verification-based studies Nie et al. (2019); Yoneda et al. (2018); Hanselowski et al. (2018); Malon (2018); Thorne et al. (2018); Dun et al. (2021); Bazmi et al. (2023) consider social media posts as a claim and search for external evidence which supports or refutes the claim made in a social media post to verify the authenticity of the post. Above evidence claims verification-based studies extract evidence from resources such as Wikipedia, knowledge graph or information available in search engines. The dynamic nature of social media conversations, frequently focused on trending or emerging topics, presents challenges in accessing current and comprehensive external information from knowledge bases or information from external sources such as news feeds. This situation often results in instances where such external resources are either unavailable or insufficiently updated to reflect the latest events or trends accurately. Consequently, in such circumstances, it's difficult to get external information which will help in verifying the authenticity of social media posts. Also, content-based fake news detection's methods in literature face challenges due to insufficient or ambiguous information in the source tweets due to limited and short text in tweets Yang et al. (2022). Similarly, social interaction network-based verification relies on extensive social data, often unavailable for all social media posts Yang et al. (2022). To overcome the above limitations with existing evidence, claim verification-based methods, content-based methods and social interaction network-based methods in literature. This study proposes *Hashtag Context-aware Fake News Detection* **HCFND**. Our proposed model **HCFND** first extracts the posts from the hashtag mentioned in the source post and relevant posts extracted from the named entity mentioned in the source post as external source information, then learns a relationship between the encoded representation of the source post and encoded representations of extracted social media posts as evidence. We conduct our experiment over two publicly available benchmark datasets. Our experimental results suggest that our proposed models outperform existing state-of-the-art methods in the literature. Observations from our experimental results suggest that our proposed model offers a notable advantage over content-based methods by not solely depending on source posts. Additionally, our proposed methods do not necessitate a large network, thus overcoming limitations associated with social interaction-based approaches. Furthermore, our approach considers information posted by other people regarding the same events as an external source, eliminating the need for knowledge-based or other external sources such as mainstream media reporting. Consequently, our method addresses the limitations inherent in existing evidence claim verification methodologies.

2. Related Work

In the literature, various studies Shu et al. (2017); Zubiaga et al. (2018); Sharma et al. (2019); Zhou & Zafarani (2020); Parikh & Atrey (2018); D'Ulizia et al. (2021) have reviewed research on fake news detection. This study conducts a retrospective examination of research specifically targeting fake news detection. As detailed in Section 1, studies on fake news detection can be grouped into three main categories: Content-based, social interaction network, and evidence claim verification. Content-based methods involve analyzing linguistic, psychological, and statistical features to assess the writing style of online documents. They concentrate on scrutinizing the textual attributes of content to identify patterns, anomalies, or indicators suggestive of fake news. Content-based methods utilize either bag-of-words features along with traditional machine learning techniques Choudhary & Arora (2021); Hakak et al. (2021); Castillo et al. (2011); Popat (2017) or neural encoding models applied to social media posts Ahn & Jeong (2019); Potthast et al. (2018); Karimi et al. (2018); Bal et al. (2020); Hamed et al. (2023); Yu et al. (2017); Jain et al. (2023). As highlighted in the study Yan et al. (2015), the sparse and limited word count in short text from social media posts significantly impacts the effectiveness of content-based methods. A social interaction network refers to the intricate web of connections through which a post spreads across social media platforms. This network encompasses various elements such as social connections, user profiles, community structures, and temporal patterns, providing a comprehensive view of information dissemination. Understanding social interaction and networks is crucial to unraveling the dynamics of information flow. Social interaction and network-based methods for fake news involve understanding social connections, user profiles, community structure and temporal patterns in the news spreading. The social interaction, network-based methods involve understanding social connections, user profiles Lu & Li (2020); Nguyen et al. (2020); Zhou et al. (2022); Cheng et al. (2021), community structure Ruchansky et al. (2017); Wang et al. (2022); Shahid et al. (2022), post comment relation Shu et al. (2019); Wu & Rao (2020); Yang et al. (2022) and temporal patterns in the tweet propagation network Shu et al. (2020) or correlation between user profiles and user interaction Lu & Li (2020); Nguyen et al. (2020); Zhou et al. (2022); Cheng et al. (2021) within tweet propagation network. Social interaction, network-based methods focus on tweet and retweet networks, but it is expensive to obtain a huge tweet propagation network for every source postYang et al. (2022). Evidence claim verification methods treat social media posts or news shared on social media as claims and then extract evidence from external sources like Wikipedia or Wikidata to support or refute these claims. Studies Nie et al. (2019); Yoneda et al. (2018); Hanselowski et al. (2018); Malon (2018); Thorne et al. (2018); Dun et al. (2021); Bazmi et al. (2023) consider social media posts as a claim, and search for external evidence which supports or refutes the claim made in a social media post to verify the authenticity of the post. Evidence claim verification methods extract evidence from sources like Wikipedia, knowledge graphs, or Google searches. However, applying these methods to fake news detection on social media presents challenges. (i) Social media is dynamic, with information spreading rapidly and evolving quickly. Fake news often spreads faster than fact-checking processes can keep up, as Wikipedia and knowledge bases may not update in real-time, leading to delays in accessing relevant information. (ii) Social media covers diverse topics, and fake news can vary widely and depend on context. Knowledge bases may lack comprehensive coverage or omit information on emerging topics, making it difficult to verify claims not well-represented in these sources. To address the limitations of existing methods in the literature, this study proposes a Hashtag Context-aware Fake News Detection (HCFND). HCFND leverages information posted under the hashtag mentioned in source tweets and relevant posts extracted from named entities mentioned in the source post instead of Wikidata and other knowledge sources.

3. Proposed Method

As mentioned in Section 1 this study proposes *Hashtag Context-aware Fake News Detection* **HCFND** for fake news detection in social media. Our proposed model, HCFND, initially extracts the named entities and hashtags mentioned in social media posts. It then retrieves posts under the corresponding hashtag and posts from the social media profile of the named entity mentioned in the source post. Subsequently, it verifies the similarity and consistency between the information presented in the source post and the extracted posts.

3.1. Retrieval of External Information

The introduction of the hashtag feature aimed to streamline users' capacity to track topics they find interesting and to spotlight trending posts Rauschnabel et al. (2019). Hashtags offer users a convenient method to delve into content pertaining to specific topics and assess the popularity of an event. They offer a way to monitor what others are discussing regarding a particular subject and the extent of excitement surrounding that event Jackson & Foucault Welles (2015); Wang et al. (2016). People with shared interests in a specific topic or theme participate in discussions by posting tweets under a common hashtag. Tweets associated with a particular hashtag center around a shared topic or event, connecting with individuals in a similar community interested in the topic Jackson & Foucault Welles (2015); Wang et al. (2016); Nam et al. (2017). Hence, hashtags provide valuable supplementary information that can enrich the content of individual source posts and also contribute to verifying the authenticity of the source post. The primary rationale for considering tweets associated with hashtags and those referencing a named entity cited in the source lies in the aim to ascertain the discourse surrounding shared events and topics. Furthermore, when tweets involve a named entity, investigating the discussions within that entity's Twitter profile offers additional perspective. Through the comparison of the source tweet with others under the same hashtag, the research aims to validate the accuracy and authenticity of the information presented in the original tweet. Given a source, post, $\mathcal{S}$ **HCFND** initially extracts a list of hashtags $\mathcal{H}_1$, $\mathcal{H}_2$... $\mathcal{H}_k$ and a list of named entities $\mathcal{E}_1$, $\mathcal{E}_2$... $\mathcal{E}_k$ present in the source tweet $\mathcal{S}$ using NLTK[1]. Next, we from the list of tweets $\mathcal{L}$ by extracting and adding all the tweets from hashtag $\mathcal{H}_\mathbf{i}$ $\forall \mathbf{i} = \mathbf{1}...\mathbf{k}$ to the list $\mathcal{L}$ and also from the list of tweets $\mathcal{U}$ by extracting and adding recent fifty tweets from the social media profile of a named entity $\mathcal{E}_\mathbf{i}$ $\forall \mathbf{i} = \mathbf{1}...\mathbf{k}$ to the list $\mathcal{U}$. Next to filter out tweets which are relevant to the source post $\mathcal{S}$ from the extracted tweet list $\mathcal{L}$ and $\mathcal{U}$ and discard the noisy tweets in the extracted tweet list $\mathcal{L}$ and $\mathcal{U}$. We utilize sentence-BERT (S-BERT) Reimers & Gurevych (2019) to encode the source post $\mathcal{S}$ and tweets in the tweet lists $\mathcal{L}$ and $\mathcal{U}$. We then estimate the cosine similarity between the encoded representation of the source post and the encoded representations of tweets in the tweet lists $\mathcal{L}$ and $\mathcal{U}$. We form a news tweet list $\mathcal{P}$ by incorporating tweets from both tweet lists $\mathcal{L}$ and $\mathcal{U}$ where the cosine similarity exceeds 0.5.

3.2. Correlation between Source Post and External Information

Figure 1 illustrates the operational diagram of the proposed **HCFND**. Given a source post $\mathcal{S}$ and a list of extracted tweets $\mathcal{P}$, we first encode $\mathcal{S}$ and each tweet $\mathcal{P}_i$ $\forall i = 1..n$ using a sentence transformer to obtain encoded representations $\mathbf{s}$ and $\mathbf{p}_i$ of $\mathcal{S}$ and tweet $\mathcal{P}_i$, where $\mathcal{P}_i$ represents the i^{th} tweet in the extracted tweet list $\mathcal{P}$. Next, we construct a matrix $\mathbf{P}$, where each row represents a post encoding, by concatenating the encoded representations $\mathbf{p_i}$ of posts, $\mathcal{P}_i$ $\forall i = 1..n$. Next, we apply multi-head attention Vaswani et al. (2017) between the encoded representation of the source post $\mathbf{s}$ and the encoded representation matrix $\mathbf{P}$ of the extracted tweet list $\mathcal{P}$. The key intuition behind applying multi-head attention between source post and extracted tweets is that if the information posted in the source post is true, then the source post will be contextually similar to the information presented in tweets of extracted tweets list $\mathcal{P}$ accordingly multi-head attention component will assign high attention weight between $\mathbf{s}$ and $\mathbf{p_i}$. Similarly, if the information posted in the source post is fake, then the source post will be less contextually similar to the information presented in tweets of extracted tweets list $\mathcal{P}$ accordingly multi-head attention component will assign low attention weight between $\mathbf{s}$ and $\mathbf{p_i}$. We define $\mathbf{s^q}$, key $\mathbf{P^k}$ and value $\mathbf{P^v}$ matrices using the Equation 1 defined below.

$$\mathbf{s^q_c}, \mathbf{P^k_c}, \mathbf{P^v_c} = \mathbf{s} \cdot \mathbf{W^q_c}, \mathbf{P} \cdot \mathbf{W^k_c}, \mathbf{P} \cdot \mathbf{W^v_c} \tag{1}$$

Where $\mathbf{W^q_c}$, $\mathbf{W^k_c}$ and $\mathbf{W^v_c}$ are learnable parameter matrices of query, key and value projections respectively, for c^{th} attention head of multi-head attention and $\cdot$ is the dot product between matrix. Subsequently, attention weigh $\mathbf{a_c}$ is defined as follows:

$$\mathbf{m} = \Big(\frac{\mathbf{s}_c^q\ (\mathbf{P}_c^k)^\top}{\sqrt{\mathbf{z}}}\Big) \tag{2}$$

$$\mathbf{a}_{c,i} = \Big(\frac{\mathbf{exp(m_i)}}{\sum_i \mathbf{exp(m_i)}}\Big) \tag{3}$$

Here $\mathbf{m}$ is the similarity vector, which indicates the similarity between the encoded representation $\mathbf{s}$ of the source post $\mathcal{S}$ and

[1]Named Entity Extraction using NLTK

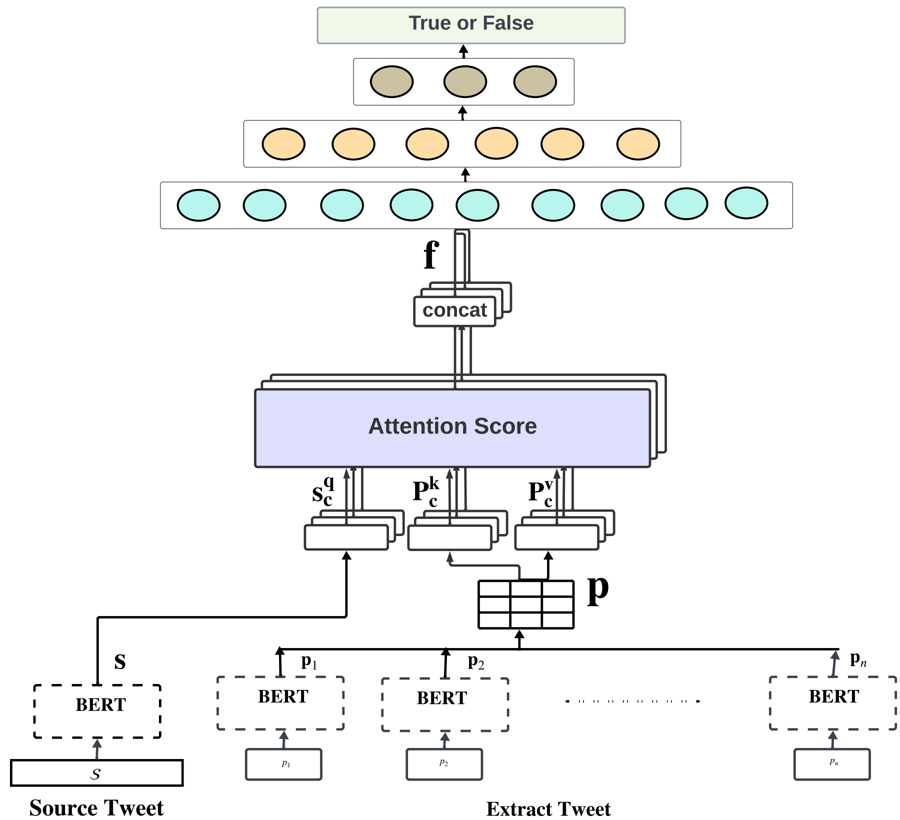

Fig. 1. Present schematic diagram of proposed model *Hashtag Context-aware Fake News Detection* **HCFND.**

the encoded representation $\mathbf{p}_i$ of i^{th} tweet in the extracted tweet list $\mathcal{P}$. Similarly, $\mathbf{a_{c,i}}$ the entry represents the similarity probability between the encoded representation $\mathbf{s}$ of the source post $\mathcal{S}$ and the encoded representation $\mathbf{p}_i$ of i^{th} tweet in the extracted tweet list $\mathcal{P}$. $\mathbf{z}$ is the dimension of $\mathbf{s_c^q}$. Next, the weighted summation is applied over the encoding of sentences $\mathbf{p_i}$ based on similarity with the encoded representation of the source post.

$$\mathbf{u_c} = \left(\mathbf{a_c}^\top \mathbf{P_c^v}\right) \tag{4}$$

Where $\mathbf{u_c}$ is the representation obtained after multiplying attention weight $\mathbf{a_c}$ and $\mathbf{P_c^v}$ which represents the encoded representation $\mathbf{p}_i$ of i^{th} tweet in the extracted tweet list $\mathcal{P}$ for $\mathbf{c^{th}}$ attention head. We obtain $\mathbf{c}$ representation $\mathbf{u_1}$, $\mathbf{u_2}$ . . . $\mathbf{u_c}$ by applying multi-head attention between the encoded representation $\mathbf{s}$ of the source post $\mathcal{S}$ and the encoded representation $\mathbf{p}_i$ of i^{th} tweet in the extracted tweet list $\mathcal{P}$. Subsequently, we obtain a representation matrix $\mathbf{U}$ by concatenating representations $\mathbf{u_i}$ obtained in different attention heads and pass to a dense layer to obtain the final representation $\mathbf{f}$.

$$\mathbf{f} = \left(\mathbf{u_1} \oplus \mathbf{u_2} \oplus .. \oplus \mathbf{u_i} \oplus . \oplus \mathbf{u_c}\right)\mathbf{W_m} \tag{5}$$

Next, we pass the final representation $\mathbf{f}$ to three layers fully connected neural network for fake news classification.

Table 1. Characteristics of Twitter15 and Twitter 16 datasets

Platform	Twitter15	Twitter16
# Source tweets	742	412
# True	372	205
# False	370	207
# Users	488,211	268,721
#Retweets per story	292.19	308.7
# Words per source	13.25	12.81

4. Experimental Setups

4.1. Dataset Characteristics

This study considers two publicly available benchmark datasets, namely Twitter15 and Twitter16 Ma et al. (2017) datasets, to assess the performance of proposed and baseline models. Table 1 presents the characteristics of the Twitter15 and Twitter16 datasets. We consider 80% for training, 10% for Validation and 10% for testing.

4.2. Experimental setups

We consider F-measure *(F.)* and Accuracy (*Acc.*) as evaluation metrics to evaluate the performance of proposed and baseline models. Table 2 presents the details of hyperparameters

Table 2. Details of hyperparameters used to produce the results presented in this paper

Hyperparameter	Values
Learning Rate	10^{-4}
Batch Size	5
Epochs	200
User Embedding Dimension	11
Reply Embedding Dimension	384
Number of Layers in Classifier	3
Number of Convolutions in GAT	3
Number of tweets from hashtag in GAT	50

Table 3. Comparison of the performances of different models over `Twitter-15` and `Twitter-16` datasets. Here, (Acc.) and (F) indicate accuracy and F-measure, respectively.

	Model		**Twitter-15**		**Twitter-16**	
			Acc.	**F1**	**Acc.**	**F1**
Baseline	**Content based**	DTC Ma et al. (2015)	0.494	0.494	0.561	0.561
		SVM-TS Ma et al. (2015)	0.519	**0.519**	0.693	**0.691**
		mGRU Kwon et al. (2017)	0.554	0.510	0.661	0.556
		RFC Kwon et al. (2017)	0.538	0.464	0.662	0.627
		tCNN Yang et al. (2018)	**0.588**	0.514	**0.737**	0.620
	Social Interaction	CSI Ruchansky et al. (2017)	0.698	0.717	0.661	0.630
		HPFN Shu et al. (2020)	0.724	0.729	0.784	**0.788**
		UPFD Dou et al. (2021)	**0.729**	**0.732**	**0.788**	0.781
	Post comment relation	PostCom2DR Yang et al. (2022)	**0.714**	**0.706**	**0.779**	**0.8**
	Knowledge Graph	MVCAN Bazmi et al. (2023)	**0.729**	**0.730**	**0.799**	**0.821**
Proposed	Model	HCFND	**0.758**	**0.762**	**0.849**	**0.856**

that have been used to produce the results presented in this paper. Though we experimented with several values of hyperparameters, Table 2 presents the values of hyperparameters for which the proposed model performance is superior.

4.3. Baselines

To evaluate the strengths and weaknesses of our proposed model, HCFND, this study conducts a comparison with recent state-of-the-art methods across different categories in the literature: (i) content-based approaches, (ii) social interaction network-based methods, (iii) post-comment relation techniques and (iv) Knowledge graph-based methods. This study considers content-based fake news detection methods The Twitter credibility model using *Decision Tree Classifier* **DTC** Ma et al. (2015), *Support Vector Machine classifier that use Time Series* **SVM-TS** Ma et al. (2015), *Gated Recurrent Unit* **mGRU** Ma et al. (2016), *Random Forest Classifier* **RFC** Kwon et al. (2017), *Capture, Score, and Integrate* **CSI** Ruchansky et al. (2017) and *Text information based Convolutinal Neural Network* **tCNN** Yang et al. (2018) as baseline models. We also consider the Knowledge graph-based method *Multi-View Co-Attention Network* **MVCAN** Bazmi et al. (2023). Apart from content-based methods and social interaction network-based methods, we also consider the post-comment interaction-based method *Post and Comments To Detect Rumors* **PostCom2DR** Yang et al. (2022) as the baseline model.

5. Results and Discussion

Table 3 compares the performance of both baseline models and the proposed **HCFND** model across the Twitter-15 and Twitter-16 datasets. The baseline models are classified into three categories: content-based methods, social interaction network-based methods, and post-comment relation methods. First, we study the response of content-based methods over Twitter-15 and Twitter-16 datasets. From Table 3, it is evident that *tCNN* Yang et al. (2018) **wrong citation** outperforms other content-based baseline methods over twitter-15 and twitter-16 datasets. However, it is also evident that the performance of content-based baseline methods over twitter-15 and twitter-16 datasets is average. This indicated that the contents in the source post alone are insufficient to determine social media posts' authenticity. Next, social interaction-based methods involve different sets of concepts. (i) CSI Ruchansky et al. (2017) model integrates post-comment interaction and user profile information to detect fake news. (ii) $HPFN$ Shu et al. (2020) mainly relies on features related to the hierarchical propagation network constructed based on tweets and retweets of source posts. (iii) $UPFD$ Dou et al. (2021) considers the relation between two hundred recent posts of users who participated in a social interaction network and source posts. From Table 3, it is evident that the social interaction-based baseline models and post-comment relation-based baseline models outperform content-based baseline models. The superior performance of social interaction-based methods and post comment relation-based baseline model establishes that considering social interaction-based information such as user profile interaction, information present in comments, and response to source posts and tweet propagation network along with social media posts improves the performance of models for fake news detection. Subsequently, we compare the performance of a Knowledge graph-based model $MVCAN$ Bazmi et al. (2023) to content-based, social interaction-based, and post-comment relation-based. Table 3 shows that the Knowledge graph-based baseline model $MVCAN$ outperforms other content-based methods, social interaction, and post-comment relation-based models over both twitter-15 and twitter-16 datasets. Such superior performance of the Knowledge graph-based method $MVCAN$ indicates that considering information related to the named entity mentioned in the source post from an external knowledge base such as Wikidata improves the performance of fake news detection models and helps efficiently verify the authenticity of social media posts. As mentioned in Section 1, the dynamic nature of social media conversations, frequently focused on trending or emerging topics, presents challenges in accessing current and comprehensive external information from knowledge bases such as Wikidata or information from external sources such as news feeds. This situation often results in such external resources needing to be more available or updated to accurately reflect the latest events or trends. The $MVCAN$ model also relies on a knowledge base from Wikidata, which may not frequently deal with social media's trending or emerging topics.

To address the drawbacks of the knowledge graph-based $MVCAN$ model, this study proposes the "Hashtag Context-aware Fake News Detection" $HCFND$ approach. $HCFND$ incorporates social media posts associated with the hashtag referenced in the source tweet and relevant posts from the social media profiles of the mentioned entities as external evidence. From Table 3, it is evident that our proposed model outperforms content-based baseline models, social interaction-based models, post-comment relation-based models, and knowledge graph-based baseline models. Consequently, we can claim that comparing the source tweet with tweets under the hashtag and relevant tweets from the named entity mentioned in the source tweet is helpful in fake news detection. Also, our proposed model overcomes the limitations of content-based methods, social interaction-based methods, post-comment relation and knowledge graph-based methods. (i) Our proposed methods go beyond solely relying on the content within source posts. Consequently, they mitigate the limitations inherent in content-based methods, such as sparsity and the limited number of words in social media posts, which significantly impact their performance. (ii)Our proposed model doesn't necessitate a large and costly social interaction network (comprising tweets and retweets) to ascertain the authenticity of social media posts. Thus, it effectively surpasses the limitation associated with social interaction-based approaches, which typically demand extensive social interaction networks to determine the authenticity of social media posts Yang et al. (2022). (iii) Since our proposed model does not depend on external databases like Wikidata, which may not be updated frequently to reflect the dynamic nature of social media trends, it relies on information sourced from other individuals within the community, namely social media posts. As a result, it effectively circumvents the constraints typically associated with knowledge graph-based methods.

6. Conclusions and Future Works

This paper proposes *Hashtag Context-aware Fake News Detection* (**HCFND**) for fake news detection over social media networks. The proposed model, denoted as HCFND, integrates data from social media by incorporating posts under a designated hashtag referenced in the source tweet and pertinent content from the social media profile associated with the mentioned entity. This external evidence is utilized to establish a connection between the encoded representation of the source post and the encoded representations of the extracted social media posts. We conducted our experiment on `Twitter-15` and `Twitter-16`. Our experimental results suggest that our proposed model outperforms existing state-of-the-art models in the literature.

References

Ahn, Y.-C., & Jeong, C.-S. (2019). Natural language contents evaluation system for detecting fake news using deep learning. In *2019 16th international joint conference on computer science and software engineering (JCSSE)* (pp. 289–292). IEEE.

Ali, I., Ayub, M. N. B., Shivakumara, P., Noor, N. F. B. M. et al. (2022). Fake news detection techniques on social media: A survey. *Wireless Communications and Mobile Computing*, *2022*.

Bal, R., Sinha, S., Dutta, S., Joshi, R., Ghosh, S., & Dutt, R. (2020). Analysing the extent of misinformation in cancer related tweets. In *Proceedings of the International AAAI Conference on Web and Social Media* (pp. 924–928). volume 14.

Bazmi, P., Asadpour, M., & Shakery, A. (2023). Multi-view co-attention network for fake news detection by modeling topic-specific user and news source credibility. *Information Processing & Management*, *60*, 103146.

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).

Cheng, L., Guo, R., Shu, K., & Liu, H. (2021). Causal understanding of fake news dissemination on social media. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining* (pp. 148–157).

Choudhary, A., & Arora, A. (2021). Linguistic feature based learning model for fake news detection and classification. *Expert Systems with Applications*, *169*, 114171.

Dou, Y., Shu, K., Xia, C., Yu, P. S., & Sun, L. (2021). User preference-aware fake news detection. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 2051–2055).

Dun, Y., Tu, K., Chen, C., Hou, C., & Yuan, X. (2021). Kan: Knowledge-aware attention network for fake news detection. In *Proceedings of the AAAI conference on artificial intelligence* (pp. 81–89). volume 35.

D'Ulizia, A., Caschera, M. C., Ferri, F., & Grifoni, P. (2021). Fake news detection: a survey of evaluation datasets. *PeerJ Computer Science*, *7*, e518.

Hakak, S., Alazab, M., Khan, S., Gadekallu, T. R., Maddikunta, P. K. R., & Khan, W. Z. (2021). An ensemble machine learning approach through effective feature extraction to classify fake news. *Future Generation Computer Systems*, *117*, 47–58.

Hamed, S. K., Ab Aziz, M. J., & Yaakub, M. R. (2023). Fake news detection model on social media by leveraging sentiment analysis of news content and emotion analysis of users' comments. *Sensors*, *23*, 1748.

Hanselowski, A., Zhang, H., Li, Z., Sorokin, D., Schiller, B., Schulz, C., & Gurevych, I. (2018). Ukp-athene: Multi-sentence textual entailment for claim verification. In *Proceedings of the First Workshop on Fact Extraction and VERification (FEVER)* (pp. 103–108).

Jackson, S. J., & Foucault Welles, B. (2015). Hijacking# mynypd: Social media dissent and networked counterpublics. *Journal of communication*, *65*, 932–952.

Jain, M. K., Gopalani, D., & Meena, Y. K. (2023). Confake: fake news identification using content based features. *Multimedia Tools and Applications*, (pp. 1–27).

Karimi, H., Roy, P., Saba-Sadiya, S., & Tang, J. (2018). Multi-source multi-class fake news detection. In *Proceedings of the 27th international conference on computational linguistics* (pp. 1546–1557).

Kwon, S., Cha, M., & Jung, K. (2017). Rumor detection over varying time windows. *PloS one*, *12*, e0168344.

Lu, Y.-J., & Li, C.-T. (2020). Gcan: Graph-aware co-attention networks for explainable fake news detection on social media. *arXiv preprint arXiv:2004.11648*, .

Ma, J., Gao, W., Mitra, P., Kwon, S., Jansen, B. J., Wong, K. F., & Cha, M. (2016). Detecting rumors from microblogs with recurrent neural networks. In *IJCAI International Joint Conference on Artificial Intelligence* (pp. 3818–3824). volume 2016.

Ma, J., Gao, W., Wei, Z., Lu, Y., & Wong, K.-F. (2015). Detect rumors using time series of social context information on microblogging websites. In *Proceedings of the 24th ACM international on conference on information and knowledge management* (pp. 1751–1754).

Ma, J., Gao, W., & Wong, K.-F. (2017). Detect rumors in microblog posts using propagation structure via kernel learning. Association for Computational Linguistics.

Malon, C. (2018). Team papelo: Transformer networks at fever. In *Proceedings of the First Workshop on Fact Extraction and VERification (FEVER)* (pp. 109–113).

Nam, H., Joshi, Y. V., & Kannan, P. (2017). Harvesting brand information from social tags. *Journal of Marketing*, *81*, 88–108.

Nguyen, V.-H., Sugiyama, K., Nakov, P., & Kan, M.-Y. (2020). Fang: Leveraging social context for fake news detection using graph representation. In *Proceedings of the 29th ACM international conference on information & knowledge management* (pp. 1165–1174).

Nie, Y., Chen, H., & Bansal, M. (2019). Combining fact extraction and verification with neural semantic matching networks. In *Proceedings of the AAAI Conference on Artificial Intelligence* (pp. 6859–6866). volume 33.

Parikh, S. B., & Atrey, P. K. (2018). Media-rich fake news detection: A survey. In *2018 IEEE conference on multimedia information processing and* *retrieval (MIPR)* (pp. 436–441). Miami, FL, USA: IEEE.

Popat, K. (2017). Assessing the credibility of claims on the web. In *Proceedings of the 26th international conference on world wide web companion* (pp. 735–739).

Potthast, M., Kiesel, J., Reinartz, K., Bevendorff, J., & Stein, B. (2018). A stylometric inquiry into hyperpartisan and fake news. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 231–240).

Rauschnabel, P. A., Sheldon, P., & Herzfeldt, E. (2019). What motivates users to hashtag on social media? *Psychology & Marketing*, *36*, 473–488.

Reimers, N., & Gurevych, I. (2019). Sentence-bert: Sentence embeddings using siamese bert-networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics.

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).

Shahid, W., Li, Y., Staples, D., Amin, G., Hakak, S., & Ghorbani, A. (2022). Are you a cyborg, bot or human?—a survey on detecting fake news spreaders. *IEEE Access*, *10*, 27069–27083.

Sharma, K., Qian, F., Jiang, H., Ruchansky, N., Zhang, M., & Liu, Y. (2019). Combating fake news: A survey on identification and mitigation techniques. *ACM Transactions on Intelligent Systems and Technology (TIST)*, *10*, 1–42.

Shu, K., Cui, L., Wang, S., Lee, D., & Liu, H. (2019). defend: Explainable fake news detection. In *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining* (pp. 395–405).

Shu, K., Mahudeswaran, D., Wang, S., & Liu, H. (2020). Hierarchical propagation networks for fake news detection: Investigation and exploitation. In *Proceedings of the International AAAI Conference on Web and Social Media* (pp. 626–637). volume 14.

Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media: A data mining perspective. *ACM SIGKDD explorations newsletter*, *19*, 22–36.

Thorne, J., Vlachos, A., Cocarascu, O., Christodoulopoulos, C., & Mittal, A. (2018). The FEVER2.0 shared task. In *Proceedings of the Second Workshop on Fact Extraction and VERification (FEVER)*.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, *30*.

Wang, H., Wang, S., & Han, Y. (2022). Detecting fake news on chinese social media based on hybrid feature fusion method. *Expert Systems with Applications*, *208*, 118111.

Wang, R., Liu, W., & Gao, S. (2016). Hashtags and information virality in networked social movement: Examining hashtag co-occurrence patterns. *Online Information Review*, *40*, 850–866.

Wu, L., & Rao, Y. (2020). Adaptive interaction fusion networks for fake news detection. In *ECAI 2020* (pp. 2220–2227). IOS Press.

Yan, R., Yen, I. E., Li, C.-T., Zhao, S., & Hu, X. (2015). Tackling the achilles heel of social networks: Influence propagation based language model smoothing. In *Proceedings of the 24th international conference on world wide web* (pp. 1318–1328).

Yang, Y., Wang, Y., Wang, L., & Meng, J. (2022). Postcom2dr: Utilizing information from posts and comments to detect rumors. *Expert Systems with Applications*, *189*, 116071.

Yang, Y., Zheng, L., Zhang, J., Cui, Q., Li, Z., & Yu, P. S. (2018). Ti-cnn: Convolutional neural networks for fake news detection. *arXiv preprint arXiv:1806.00749*, .

Yoneda, T., Mitchell, J., Welbl, J., Stenetorp, P., & Riedel, S. (2018). Ucl machine reading group: Four factor framework for fact finding (hexaf). In *Proceedings of the First Workshop on Fact Extraction and VERification (FEVER)* (pp. 97–102).

Yu, F., Liu, Q., Wu, S., Wang, L., Tan, T. et al. (2017). A convolutional approach for misinformation identification. In *IJCAI* (pp. 3901–3907).

Zhou, X., & Zafarani, R. (2020). A survey of fake news: Fundamental theories, detection methods, and opportunities. *ACM Computing Surveys (CSUR)*, *53*, 1–40.

Zhou, X., Zafarani, R., & Ferrara, E. (2022). From fake news to# fakenews: Mining direct and indirect relationships among hashtags for fake news detection. *arXiv preprint arXiv:2211.11113*, .

Zubiaga, A., Aker, A., Bontcheva, K., Liakata, M., & Procter, R. (2018). Detection and resolution of rumours in social media: A survey. *ACM Computing Surveys (CSUR)*, *51*, 1–36.