A Multi-classification Division-aggregation Framework for Fake News Detection

This article has been accepted for publication in IEEE Transactions on Big Data. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TBDATA.2024.3378098

A Multi-classification Division-aggregation Framework for Fake News Detection

Wen Zhang, Haitao Fu, Huan Wang*, Zhiguo Gong, *Senior Member*, *IEEE*, Pan Zhou, and Di Wang

***Abstract*—Nowadays, as human activities are shifting to social media, fake news detection has been a crucial problem. Existing methods ignore the classification difference in online news and cannot take full advantage of multi-classification knowledges. For example, when coping with a post "A mouse is frightened by a cat," a model that learns "computer" knowledge tends to misunderstand "mouse" and give a fake label, but a model that learns "animal" knowledge tends to give a true label. Therefore, this research proposes a multi-classification division-aggregation framework to detect fake news, named CKA, which innovatively learns classification knowledges during training stages and aggregates them during prediction stages. It consists of three main components: a news characterizer, an ensemble coordinator, and a truth predictor. The news characterizer is responsible for extracting news features and obtaining news classifications. Cooperating with the news characterizer, the ensemble coordinator generates classification-specifical models for the maximum reservation of classification knowledges during the training stage, where each classification-specifical model maximizes the detection performance of fake news on corresponding news classifications. Further, to aggregate the classification knowledges during the prediction stage, the truth predictor uses the truth discovery technology to aggregate the predictions from different classification-specifical models based on reliability evaluation of classification-specifical models. Extensive experiments prove that our proposed CKA outperforms state-of-the-art baselines in fake news detection.**

***Index Terms*—Fake news detection, classification-specifical model, truth discovery, multi-classification knowledge**

I. INTRODUCTION

WITH the escalated increase of human activities in the world across online services, more and more people tend to seek out real-time news from social media platforms rather than traditional news organizations [1][2]. As people continue to benefit from the convenience of accessing social media, they also expose themselves to a large amount of fake news. Since the negative influence of fake political news captures the world's attention in the 2016 U.S. presidential election, fake news detection has been a challenging research area for several years. The spread of fake news has continually hit new highs with hot topics triggering extensive distribution and generating aggressive opinions [3]. Therefore, it is important to develop an effective method to detect fake news to restrict their negative influence.

Thus far, existing deep learning methods have achieved impressive performances due to the strong ability of feature extraction [4]. However, existing methods ignore the potential classification difference in online news and cannot take full advantage of multi-classification knowledges. Because of creation uncertainty and annotation difficulty, it is time-consuming to collect sufficient verified news in one classification and train the model for unverified news in the same classification [5]. Available training datasets usually collect the verified news from different news classifications and contain multi-classification knowledges. The news items from each news classification have similar word frequencies and propagation patterns [6], and the deep learning model trained on each news classification tend to learn the specifical classification knowledge to characterize the corresponding news classification. To avoid the waste of classification knowledges and maximize the use of verified news from different news classifications, exploiting multi-classification knowledges to improve the detection performance of fake news is of great importance.

This research considers the classification difference in online news and attempts to take full advantage of multi-classification knowledges in fake news. *The first issue required to be addressed is how to learn classification knowledges based on verified news from different news classifications during the training stage?* The verified news items from different news classifications have their specifical classification knowledges to characterize the training processes of deep learning models. It is challenging to learn specifical classification knowledges in an independent manner due to the possible lack of sufficient verified news on each classification, especially for newly emerged classifications. Inspired by ensemble learning that strategically creates and combines multiple models to seek better predictive performances [7], we generate different classification-specifical models for different news classifications to learn their specifical classification knowledges in a collaborative manner.

Wen Zhang is with the College of Informatics, Huazhong Agricultural University, Wuhan 430070, China. E-mail: zhangwen@mail.hzau.edu.cn.

Haitao Fu is with the College of Informatics, Huazhong Agricultural University, Wuhan 430070, China. E-mail: fuhaitao@webmail.hzau.edu.cn.

Huan Wang is with the State Key Laboratory of Internet of Things for Smart City, Department of Computer Information Science, University of Macau, Macau, China. E-mail: hwang@mail.hzau.edu.cn.

Zhiguo Gong is with the State Key Laboratory of Internet of Things for Smart City, Department of Computer Information Science, University of Macau, Macau, China. E-mail: fstzgg@um.edu.mo.

Pan Zhou is with the School of Cyber Science and Engineering, Huazhong University and Science and Technology, Wuhan 430074, China. E-mail: panzhou@hust.edu.cn.

Di Wang is with the Division of Computer, Electrical and Mathematical Sciences and Engineering at the King Abdullah University of Science and Technology, Thuwal, Saudi Arabia. E-mail: di.wang@kaust.edu.sa.

During the training stage, the construction of different classification-specifical models drives the division process of classification knowledges. *The second issue is how to aggregate different classification knowledges to predict the fake likelihoods of unverified news during the prediction stage?* Each unverified news item may belong to a historical or newly emerged classification. It is difficult to maintain the classification consistency between unverified news and learned classification knowledges. Inspired by truth discovery that overcomes the conflicts among the collected multi-source information [8], we consider each generated classification-specifical model as a source and integrate multi-source information by estimating the reliability of each source. The predictions from different classification-specifical models are aggregated to represent the aggregation of classification knowledges in fake news detection.

To tackle the issues above, we propose a multi-classification division-aggregation framework in this research, named CKA. It innovatively learns classification knowledges during training stages and aggregates them during prediction stages. It consists of a news characterizer, an ensemble coordinator, and a truth predictor. First, the news characterizer extracts news features and applies TwitterNews+ [9] to divide news into different news classifications. Second, inspired by ensemble learning, the ensemble coordinator generates different classification-specifical models to learn specifical classification knowledges, and each classification-specifical model maximizes the detection performance of fake news on corresponding news classifications. Third, the truth predictor applies truth discovery technologies to estimate the reliability of each classification-specifical model to predict the fake likelihoods of unverified news. The main contributions of this research are summarized as follows:

- We recognize the importance of multi-classification knowledges in fake news detection and propose a framework to exploit multi-classification knowledges during training and prediction stages to improve the detection performance of fake news.
- During the training stage, the model coordinator is proposed to generate classification-specifical models for different news classifications, and each classification-specifical model tends to learn specifical classification knowledges by maximizing the detection performance of fake news on its corresponding news classification.
- During the prediction stage, the truth predictor is proposed to apply the truth discovery technology to aggregate classification knowledges from different classification-specifical models to detect fake news.
- Our proposed CKA is a general framework to exploit multi-classification knowledges in fake news detection. The integrated news feature extraction in the news characterizer can be readily replaced with existing modules specifically designed to extract multi-modal features from news.

The remainder of this paper is structured as follows. In Section II, we summarize the related literature survey. The problem of fake news detection is defined in Section III. In Section IV, we detail the proposed CKA. In Section V, we conduct extensive experiments to verify the performance. Section VI concludes this research and outlines future directions.

II. Background

In this section, we briefly introduce preliminary concepts, including ensemble learning and truth discovery, then review work related to our research.

A. Preliminary Concepts

Ensemble Learning. Ensemble learning is a technique used to improve the generalization performance of tasks such as classification, prediction, and function approximation by combining multiple independent models. The idea behind ensemble learning is that by combining several weak or mediocre models, a more accurate and reliable prediction can be made. Deep ensemble learning combines the advantages of deep learning models and ensemble learning, resulting in improved generalization performance of the final model [10]. The main goal of ensemble learning is to reduce the likelihood of unintentionally selecting poor models and to improve the final performance. Various strategies are used for ensemble learning, including assigning reliability to model decision points, selecting optimal features, data fusion, incremental learning, non-stationary learning, and error correction [11].

Truth Discovery. Truth discovery is a process of determining the most accurate and reliable information from a set of conflicting sources. As the amount of available information continues to increase, interested object data can be collected from more and more sources. However, for the same object, there are often conflicts between the collected multi-source information [12]. To address this challenge, the research line of truth discovery has become a hot topic, which integrates multi-source information by estimating the reliability of each source [13]. The process involves two steps: source reliability estimation and truth discovery. In truth discovery, there is no supervision, and the reliability of each source can only be inferred from the multi-source information. The principle used is that sources that provide true information more frequently will be assigned higher reliability degrees, and information supported by reliable sources will be regarded as truths [14].

B. Related Work

Due to the proliferation of fake news on social media in recent years, research on fake news detection has attracted the attention of many researchers [15,16].

Several traditional machine learning methods have been employed to detect fake news, such as Ahmed et al.'s proposal of a fake news detection model using n-gram analysis and machine learning techniques [17]. They compared two feature extraction techniques and six machine classification techniques. Gravanis et al. [18] used content-based features and machine learning algorithms for fake news detection based on contextual features. They tested various popular machine learning models and analyzed improvements using ensemble machine learning strategies like AdaBoost and Bagging. Additionally, Conroy et al.

This article has been accepted for publication in IEEE Transactions on Big Data. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TBDATA.2024.3378098

[19] combined semantic analysis and machine learning to propose an automatic detection model. The model used a support vector machine for classification by obtaining keyword information and word frequencies. While these machine learning methods relied on artificial features extracted from news text or structure, manual extraction of such features was labor-intensive, limiting their performance and utility [20].

Deep learning methods can extract online news features more effectively in fake news detection than traditional machine learning methods. For instance, Liu and Wu [21] proposed a deep neural network for early detection of fake news. The network consisted of a state-sensitive crowd response feature extractor, a location-aware attention mechanism, and a multi-region mean pooling mechanism. Ma et al. [22] proposed a recurrent neural network-based method for rumor recognition that learned the continuous representation of microblog events. Shu et al. [23] demonstrated how to effectively derive and utilize weak supervision for learning in situations with limited labeled data, proposing a new type of weak supervision to detect fake news from multifaceted social media data. Song et al. [24] developed a fake news detection method based on dynamic propagation graph that models each news dissemination graph as a series of graph snapshots recorded in discrete time steps. Davoudi et al. [25] used recursive neural networks to encode the evolution patterns of propagation trees and stance networks over time, aggregating the outputs of these components to determine the authenticity of news articles. Meel et al. [26] proposed a convolutional neural network semi-supervised framework that utilizes the linguistic and stylistic information of annotated news articles to explore hidden patterns in unlabeled data. The framework is based on the concept of self-integration, which allows for innovative and effective detection of fake news. In addition, Wu et al. [27] proposed a Category-controlled Encoder-Decoder model to enhance the detection of fake news by considering the category of a news as true or real. This model generated examples with category-differentiated features, thereby extending the dataset capacity and achieving a data enhancement effect. Subsequently, Li et al. [28] combined metric learning and frequency analysis for improved face forgery detection, aiding fake news identification through nuanced discrepancies. Shang et al. [29] introduced a Spatio-Temporal Graph Network that uses spatial-temporal graphs to detect video face forgeries through inconsistencies, enhancing fake news detection by analyzing spatial and temporal data.

However, the research above ignores the role of multi-classification knowledges in fake news detection, which restricts the maximum advantage of domain knowledge from different news classifications. Therefore, this research attempts to exploit multi-classification knowledges to improve the performance in fake news detection.

III. Problem Definition

The frequently used notations in this research are summarized in Table 1. With the popular of the online social media, a large amount of fake news can spread quickly and easily, making it difficult for individuals to distinguish between what is true and what is false. In this research, we use N^l and Y^l to denote the sets of verified news and corresponding labels, respectively. N^u denotes the set of unverified news. Formally, we define the fake news detection problem in this research as follows. Given N^l, Y^l, and N^u, we need to obtain a fake news set $A = \{n | n \in N^u \wedge n\ is\ fake\ news\}$, where the fake ones in N^u is detected. Detecting fake news involves determining whether unverified news is true or fake, essentially a binary classification problem.

TABLE 1
FREQUENTLY-USED SYMBOLS IN THIS RESEARCH

Symbol	Meaning	Symbol	Meaning
N^l	Set of verified news.	Y^l	Label set of news in N^l.
N^u	Set of unverified news.	n	A news item.
K	Classification Number.	∂	A classification-specifical model.
N_k	Set of news in the k-th news classification.	e^{∂}	Reliability of the classification-specifical model ∂.
$P(n)$	Fake likelihood of news n.	x_i	Embedding vector of i-th word.
Ψ	Set of classification-specifical models.	V	News label set.
θ_f	Learned parameters in the news characterizer.	$\theta_d^{(k)}$	Set of parameters in the kth classification-specifical model.
E_n	Feature vector of news n.	v_n	Candidate label.
M	Matching set.	I	Word number.
r_i	Reset gate.	z_i	Update gate.
$P^{(k)}(n)$	Fake likelihood of the news n by the k-th classification-specifical model.	m	Number of possible window sizes.

IV. Methodology

In this section, we discuss in detail the proposed CKA. The architecture of CKA is illustrated in Section A. The news characterizer, the ensemble coordinator, and the truth predictor are three key components of CKA. Section B explains the news characterizer to extract news features and obtain news classifications. Section C details the ensemble coordinator to generate classification-specifical models to detect fake news. Section D details the truth predictor to aggregate the predictions from different classification-specifical models. Section E illustrates the integrated CKA to detect fake news.

A. Framework Outline

The architecture of CKA is shown in Fig. 1, where the news characterizer, the ensemble coordinator, and the truth predictor cooperate in exploiting multi-classification knowledges to improve the performance in fake news detection. First, the news characterizer extracts news features and obtains news classifications. Cooperating with the news characterizer, the ensemble coordinator simultaneously generates classification- specifical models for different news classifications and reduces the correlations among classification-specifical models. In the ensemble

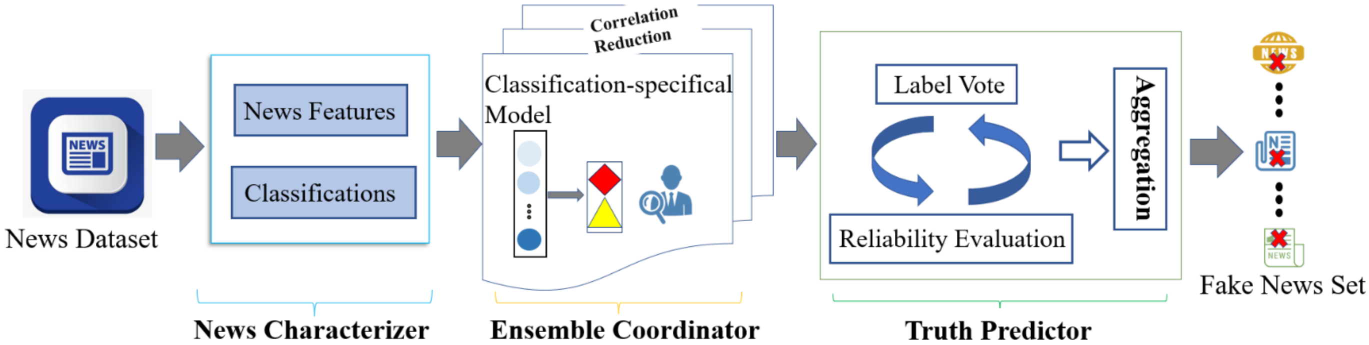


Fig. 1. The architecture of our proposed CKA.

coordinator, we believe that the maximum reservation of classification knowledge is beneficial to weaken knowledge conflicts among news classifications. Finally, the truth predictor is proposed to aggregate the predictions from different classification-specifical models based on reliability evaluation of classification-specifical models.

B News Characterizer

The news characterizer is introduced to extract news features and obtain news classifications.

News Feature Extraction. The news characterizer uses CNN and TwitterNews+ [9] as the core calculation modules for extracting news features and obtaining news classifications. Specifically, the module of CNN can meet our needs in this research, and other state-of-the-art models designed for single-modal or multi-modal feature extractions can be easily used to replace it. CNN has been widely used in many actual applications [30], where the embedding vector for each word is initialized with the pre-trained word embedding by word2vec [31]. We use I to represent the maximum word length of news. For the i-th word in the news item n, the d-dimensional word embedding vector is represented as x_i ($i \in [1, I]$). The news item n with I words is represented as:

$$x_{1:I} = x_1 \oplus x_2 \oplus \ldots \oplus x_i \oplus \ldots \oplus x_I \tag{1}$$

Here, $\oplus$ is the concatenation operator. Based on the input of the contiguous sequence of s words in a news item, a convolutional filter with the window size s outputs one feature. For the contiguous sequence of s words starting with the ith word, the filter operation is as follows:

$$X_i = \sigma(W_c \cdot x_{i:i+s-1}) \tag{2}$$

Here, $\sigma(\cdot)$ denotes the leaky ReLU activation function and W_c is the weight of this filter. The feature vector for the news item n is calculated as:

$$X = [X_1, X_2, \ldots, X_i, \ldots X_{I-s+1}] \tag{3}$$

For every X, the max-pooling operation with the maximum value is used to extract the most important textual information and the corresponding feature from each filter is obtained. This process is repeated until the features from all filters are obtained. Assuming that there are m possible window sizes and each window size owns h filters, we have $m * h$ filters to learn $m * h$ features for each news item. After the max-pooling layer, the textual features can be represented as $R_{x_c} \in \mathbb{R}^{m \cdot h}$. We use W_{cf} to represent the weight matrix of the fully connected layer. Following the max-pooling operation, we use a fully connected layer to ensure the feature representation $R_C \in \mathbb{R}^p$ in the CNN layer through the following operation:

$$R_C = \sigma\left(W_{cf} \cdot R_{x_c}\right) \tag{4}$$

We represent the news characterizer as $G_f(N; \theta_f)$, where N is the set of inputted news and θ_f is the set of learned parameters.

News Classification Obtaining. Considering the classification difference in news datasets, we introduce K to denote the number of news classifications. TwitterNews+ [9] offers a cost-effective solution for detecting news classifications in online news by utilizing specialized inverted indices and an incremental clustering approach. By utilizing random indexing-based term vectors with locality sensitive hashing, it is able to make quick decisions on the novelty of incoming news items. Also, the incremental clustering-based approach makes it efficient to cluster classification related news. The news characterizer applies TwitterNews+ [9] to divide verified news in N^l into different news classifications and obtain the optimal K value. We use f to represent the mapping function from the news in N^l to the K news classifications, where $f(n)$ denotes the classification of the news $n \in N^l$.

To address potential consequences of incorrect categorization, we propose utilizing multiple existing news categorization methods instead of relying solely on TwitterNews+ [9]. By leveraging the shared news items at each categorization across different methods, we can improve the accuracy of news categorization. Additionally, we recommend using news items with high correction rates for each categorization to generate classification-specifical models, ensuring the accuracy of news categorization.

C Ensemble Coordinator

The ensemble coordinator generates classification-specifical models for the maximum reservation of classification knowledges during the training stage. It consists of two main steps: generation for classification-specifical models and correlation reduction among classification-specifical models.

Generation for Classification-specifical Models. The ensemble coordinator is responsible for generating K classification-specifical models based on the obtained news classifications. We represent the k-th classification-specifical model as $G_d^{(k)}(\cdot; \theta_d^{(k)})$, where $\theta_d^{(k)}$ represents the set of parameters for the k-th classification-specifical model, and $k \in [1, K]$. For a given news $n \in N$, the k-th classification-specifical model evaluates its likelihood to be a fake one through the following operation:

$$P^{(k)}(n) = G_d^{(k)}(G_f(\{n\}; \theta_f^{(k)}); \theta_d^{(k)}) \tag{5}$$

$\theta_f^{(k)}$ represents the specifical θ_f for the k-th classification-specifical model. In accordance with $P^{(k)}(n)$, we introduce the modified cross entropy $O^{(k)}(n)$ for the news n as:

$$O^{(k)}(n) = -[\, y_i \, log\left(P^{(k)}(n)\right) + (1 - y_i) log\, (1 - P^{(k)}(n))] \tag{6}$$

Here, $y_i \in \{0,1\}$. When the news n is fake, $y_i = 1$. Otherwise, $y_i = 0$. A higher value of $O^{(k)}(n)$ indicates a better detection performance for the news n by the k-th classification-specifical model. The detection loss function to measure the detection performance is defined as:

$$L^{(k)}\left(\theta_f^{(k)}, \theta_d^{(k)}\right) = \frac{1}{|N|}\left[\sum_{n \in N_k} \alpha\, O^{(k)}(n) + \sum_{n \in N_k^-} (1-\alpha) O^{(k)}(n)\right] \tag{7}$$

N_k represents the set of news in the k-th classification, and the set of news in the remaining news classifications is denoted as $N_k^- = \{n | n \notin N_k \land n \in N\}$. α ($\alpha \in [0, 1]$) is introduced for the k-th classification-specifical model to balance the detection performances in the k-th classification and the other news classifications. Each classification-specifical model aims to maximize the detection performance of fake news within its specifical classification. The k-th classification-specifical model is controlled to minimize the $L^{(k)}\left(\theta_f^{(k)}, \theta_d^{(k)}\right)$ value to achieve a good detection performance.

Correlation Reduction among Classification-specifical models. All classification-specifical models need to reduce their classification correlations in fake news detection and cooperate to obtain their specifical classification knowledge. For the news n, the correlation loss between the k_1-th classification-specifical model and the k_2-th classification-specifical model is calculated by modified cross entropy as:

$$Q^{(k_1,k_2)}(n) = -[P^{(k_1)}(n) log P^{(k_2)}(n) + \left(1 - P^{(k_1)}(n)\right) log\left(1 - P^{(k_2)}(n)\right)] \tag{8}$$

When the $Q^{(k_1,k_2)}(n)$ value is larger, it indicates that the classification knowledge of the k_1-th and k_2-th classification-specifical models are more specifical. The averaged correlation loss among all classification- specifical models for all news in the set N is then defined as:

$$L_c\left(\theta_f^{(*)}, \theta_d^{(*)}\right) = \frac{2}{K(K-1)} \sum_{k_1=1}^{K} \sum_{k_2=k_1+1}^{K} \sum_{n \in N} Q^{(k_1,k_2)}(n) \tag{9}$$

To minimize the detection loss of each classification-specifical model and maximize the averaged correlation loss among all classification-specifical models, we define the comprehensive loss function as:

$$L_a\left(\theta_f^{(*)}, \theta_d^{(*)}\right) = \frac{1}{K} \sum_{k=1}^{K} L^{(k)}\left(\theta_f^{(k)}, \theta_d^{(k)}\right) - \lambda L_c\left(\theta_f^{(*)}, \theta_d^{(*)}\right) \tag{10}$$

Here, λ denotes a trade-off parameter between detection loss and correlation loss. The news characterizer $G_f(N; \theta_f)$ needs to cooperate with each classification-specifical model $G_d^{(k)}(\cdot;\, \theta_d^{(k)})$ to minimize the comprehensive loss $L_a\left(\theta_f^{(*)}, \theta_d^{(*)}\right)$ by seeking the optimal parameters of $\theta_f^{(*)}$ and $\theta_d^{(*)}$.

D Truth Predictor

Based on generated classification-specifical models from the ensemble coordinator, the truth predictor is designed for their reliability evaluation and prediction aggregation.

Reliability Evaluation of Classification-specifical models. We first evaluate the reliability of each classification-specifical model via an iterative truth discovery manner [32], consisting of a label vote process and a reliability evaluation process. These two interdependent processes are iteratively conducted until convergence.

(1) In the label vote process, the reliability of each classification-specifical model is assumed to be fixed. The label of news can be inferred through aggregating reliability votes by classification-specifical models. We use V to denote the set of possible news labels, where $V = \{fake, real\}$. For each unverified news n, we represent its current label as v_n, and $v_n \in V$. Each candidate label v_n receives the vote score $s(v_n)$ from the classification-specifical models in the following operation:

$$s(v_n) = \left(\sum_{\partial \in \Psi_{v_n}} \frac{e^{\partial}}{|N^{\partial}|}\right)^{\mu} \tag{11}$$

Here Ψ_{v_n} represents the set of classification-specifical models that provide the news label v_n for the unverified news n. N^{∂} is the number of unverified news detected by the classification-specifical model ∂. e^{∂} denotes the reliability of the classification-specifical model ∂. μ is an adjustment parameter used to maintain flexibility. The label vote process for news follows the principle that the news label from the reliable classification-specifical models will have a higher weight in the aggregation. A news label receives its vote score based on the reliabilities of all classification-specifical models that provide that news label.

(2) In the reliability evaluation process, unlike the label voting process where classification-specifical models invest their reliabilities among provided news labels, the classification-specifical models collect proportional reliabilities back from the news labels they have provided, as follows:

$$e^{\partial} = \sum_{n \in N} \sum_{v \in V} \left(s(v_n) \cdot \frac{e^{\partial} / |N^{\partial}|}{\sum_{\partial' \in \Psi_{v_n}} \frac{e^{\partial'}}{|N^{\partial'}|}} \right) \tag{12}$$

Here the classification-specifical model that provides reliable news labels more often will get more reliabilities back and have higher reliabilities. The iterative application of e^{∂} in successive computations for Equations (11) and (12) promotes convergence of the model reliabilities over iterations. The term e^{∂} ultimately denotes the converged reliability value for each model.

Prediction Aggregation of Classification-specifical models. Since each classification-specifical model has specifical classification knowledge, we enhance the importance of classification-specifical models in fake news detection of their familiar news classifications. According to the news characterizer, we obtain the feature vectors of the known verified news in the

k-th classification and calculate their averaged feature vector $E^{(k)}$. After obtaining the feature vector E_n of the unverified news n by the news characterizer, we define the consistence value $r_n^{(k)}$ as:

$$r_n^{(k)} = \frac{E^{(k)} \cdot E_n}{\|E^{(k)}\| \times \|E_n\|} \tag{13}$$

Here $r_n^{(k)}$ quantifies the consistence between the news n and the k-th classification based on cosine similarity, which focuses on the vector difference between E_n and $E^{(k)}$. Assuming that the k-th classification-specifical model predicts the likelihood of the news n to be fake is $P^{(k)}(n)$, we can obtain the final likelihood of the news n to be fake as:

$$P(n) = \sum_{\partial \in \Psi} e^{\partial} \cdot r_n^{(k)} \cdot P^{(k)}(n) \tag{14}$$

Here, Ψ represents the set of all classification-specifical model. $P(n)$ is constructed on the prediction aggregation of all classification-specifical models.

E Framework Integration

By integrating our news characterizer, ensemble coordinator, and truth predictor, the proposed CKA aims to reduce knowledge conflicts among news classifications and enhance the detection performance of fake news. Initially, we utilize the news characterizer $G_f(N;\theta_f)$ to learn the representations R_C of the news in the N^l. The news characterizer $G_f(N;\theta_f)$ leverages TwitterNews+ [9] to categorize verified news in N^l into different news classifications and determine the optimal K value. Subsequently, based on the news representations R_C from the news characterizer $G_f(N;\theta_f)$, we input the training set $\{N^l, Y^l\}$ into the ensemble coordinator to generate each classification-specific model $G_d^{(k)}(\cdot;\, \theta_d^{(k)})$ for each classification. In the ensemble coordinator, the parameter set we seek corresponds to the saddle point of the comprehensive loss function in Equation (10). Here, we denote the learning rate as η. Through stochastic gradient descent, we update the θ_f^* and θ_d^* according to Equations (15) and (16).

$$\theta_f^{(*)} \leftarrow \theta_f^{(*)} - \eta \frac{\partial L_a}{\partial \theta_f^{(*)}} \tag{15}$$

$$\theta_d^{(*)} \leftarrow \theta_d^{(*)} - \eta \frac{\partial L_a}{\partial \theta_d^{(*)}} \tag{16}$$

Finally, each classification-specific model $G_d^{(k)}(\cdot;\, \theta_d^{(k)})$ is employed to calculate the $P^{(k)}(n)$ for each unverified news n in the set of unverified news N^u. Subsequently, the truth predictor combines the $P^{(k)}(n)$ from all classification-specific models using Equation (14) to obtain the final likelihood $P(n)$. We assess the fake likelihood of each unverified news and generate an output matching set M. The process of the proposed CKA can be summarized as follows.

Method: CKA

Input: N^l— Set of verified news.
Y^l—Set of the labels of the news in N^l.
N^u— Targeted set of unverified news.

Output: M—Matching set.

```
1: M = ∅.
2: Use TwitterNews+ to obtain the number of news classifications.
3: For each training iteration do
4:     For each k-th classification do
5:         Update parameters in the news characterizer as:
6:             θ_f^(k) ← θ_f^(k) − η ∂L_a/∂θ_f^(k)
7:         Update parameters in the ensemble coordinator as:
8:             θ_d^(k) ← θ_d^(k) − η ∂L_a/∂θ_d^(k)
9:     End for
10: End for
10: For each unverified news n in N^u do
11:     Calculate P(n) by Equation (14).
12:     M = M + {(n, P(n))}
13: End for
13: Output M
```

V. Experiments

This section describes an extensive experimental study. Section A introduces the experimental datasets. In Section B, we introduce the comparison methods. Section C clarifies the details of the implementation. In Section D, we compare the CKA with state-of-the-art methods to verify its performance. As an important parameter in CKA, the setting of the classification number K is discussed in Section E. Except for the replaceable module of the news characterizer in the CKA, the importance of the ensemble coordinator and the truth predictor are analyzed in Sections F and G, respectively.

A Datasets

Our experiments involve four real-world datasets. The first dataset is OFFSY [33], which contains objective ground-truth labels and is utilized for detecting fake news from Weibo. All the verified fake news is crawled from May 2012 to January 2016 on the official fake debunking system of Weibo, which acts as the authoritative source of fake news according to prior research. This dataset comprises of 4749 fake news and 4779 true news. The second dataset is BIEN [34], which is released by the competition named Internet Fake news detection Challenge from Weibo. It contains 16602 fake news and 16365 true news. Subsequantly, the Weibo21 [6] dataset is used, which has classification label annotation. It consists of 4488 fake news and 4640 real news from nine different news domains, namely Science, Military, Education, Disasters, Politics, Health, Finance, Entertainment, and Society. Each domain can be considered as a classification. Lastly, the GossipCop dataset contains 7974 real news samples and 2036 fake news samples for training, as well as 2285 real news samples and 545 fake news samples for testing [35].

At the beginning of the experiment, these datasets are preprocessed, including removing stop words and spaces from the text data. For these datasets, we divide them into training, validation, and testing sets in an 8:1:1 ratio. Subsequently, we utilize TwitterNews+ [9] to categorize the news within the training set into different news classifications, with each news item in the training set being assigned a specific news classification.

B Baselines

To analyze the performance of CKA, we use deep learning models and machine learning models as comparison baselines.

This article has been accepted for publication in IEEE Transactions on Big Data. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TBDATA.2024.3378098

- ***FNED*** [21]: First, a status-sensitive crowd response feature extractor extracts both text features and user features from combinations of users' text response and their corresponding user profiles. Second, a position-aware attention mechanism highlights important user responses at specifical ranking positions. Third, a multi-region mean-pooling mechanism aggregates features based on multiple window sizes.
- ***MRNN*** [22]: MRNN is a fake news detection method that leverages a multilayer GRU network to model the microblog as a variable-length time series, which has proven to be effective. The approach obtains latent representations by averaging the outputs of GRU at the final time step. These representations are then fed into a fully connected layer that uses leaky ReLU as an activation function. The LSTM hidden size is set to 48.
- ***LSE*** [23]. LSE is a fake news detection method that utilizes a single layer in the long short-term memory to extract text features, while sentence embeddings are obtained from the content credibility evaluator. The approach employs LSTMs with hidden layers of size 256, and a fully connected layer that takes the extracted features as input to output the likelihood of each news being fake.
- ***DPG*** [24]: DPG is a fake news detection method based on dynamic propagation graphs that captures missing dynamic propagation information in static networks to classify fake news. The approach models each news propagation graph as a series of graph snapshots that are recorded at discrete time steps.
- ***INNSF*** [25]: INNSF is a semi-supervised Convolutional Neural Network framework that utilizes the self-ensembling concept to leverage both linguistic and stylometric information from annotated news articles, as well as discover hidden patterns in unlabeled data. By accumulating ensemble predictions, the framework aims to improve the predictive accuracy for unknown labels compared to the most recent training epoch output.

C Implementation Details

In the news characterizer, we set $k = 128$ for dimensions of word-embedding. $n_h = 16$ and the window sizes of filters are 1, 3, and 5 in Text-CNN [36]. We set the out channel $n_h = 16$ and the training epoch is 20. The iteration number is 100 and $\mu = 1.1$. We set $\alpha = 2/3$ to balance the role of the classification-specifical model in its specifical classification and the other news classifications. We set $\lambda = 0.1$ without tuning the trade-off parameter. To seek optimal parameters, we use Adam

TABLE 2
PERFORMANCES OF CKA AND COMPARISON METHODS

Dataset	*Method*	*Accuracy*	*Precision*	*Recall*	F_1
OFFSY	*FNED*	0.756	0.764	0.754	0.734
	MRNN	0.813	0.813	0.798	0.812
	LSE	0.824	0.845	0.813	0.803
	DPG	0.795	0.746	0.763	0.798
	INNSF	0.812	0.721	0.678	0.619
	CKA	**0.878**	**0.890**	**0.856**	**0.871**
BIEN	*FNED*	0.916	0.925	0.895	0.932
	MRNN	0.931	0.913	0.902	0.918
	LSE	0.893	0.812	0.905	0.922
	DPG	0.876	0.889	0.796	0.739
	INNSF	0.912	0.912	0.851	0.672
	CKA	**0.986**	**0.959**	**0.964**	**0.976**
Weibo21	*FNED*	0.767	0.729	0.743	0.802
	MRNN	0.782	0.812	0.659	0.815
	LSE	0.633	0.721	0.723	0.788
	DPG	0.657	0.803	0.789	0.718
	INNSF	0.693	0.723	0.834	0.783
	CKA	**0.794**	**0.879**	**0.915**	**0.886**
GossipCop	*FNED*	0.725	0.687	0.798	0.784
	MRNN	0.664	0.734	0.677	0.824
	LSE	0.648	0.683	0.857	0.745
	DPG	0.712	0.861	0.845	0.784
	INNSF	0.693	0.834	0.868	**0.834**
	CKA	**0.742**	**0.894**	**0.956**	0.823

This article has been accepted for publication in IEEE Transactions on Big Data. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TBDATA.2024.3378098

[37] as the optimizer. In order to stabilize the training process, we set $\eta = 0.001$. For the datasets of OFFSY, BIEN, and GossipCop, we use the TwitterNews+ [9] to automatically divide the news in the training set into different news classifications as prior information, and each verified news is assigned with one optimal classification. Based on the implementation of TwitterNews+ [9], $K = 5$, $K = 3$, and $K = 9$ are calculated for OFFSY, BIEN, and GossipCop as their optimal classification numbers, respectively. $K = 9$ is set for the dataset of Weibo21 according to its original classification information. The corresponding parameter settings of comparison methods are set according to their original references. We train the CNN from scratch.

D Performance Comparison

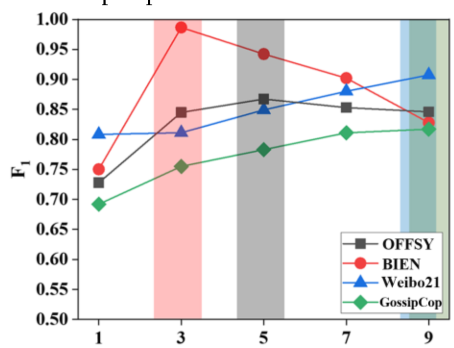
Extensive comparison experiments are implemented to analyze the performance of our proposed CKA. Based on the datasets of OFFSY, BIEN, Weibo21, and GossipCop, we evaluate the detection performances of CKA and comparison methods. Their detection performances evaluated by the metrics $Accuracy$, $Precision$, $Recall$, and F_1 [21, 22, 25] are shown in Table 2. The largest values are emphasized in bold.

Among six methods, the CKA achieves the best detection performances of fake news. The $Accuracy$, $Precision$, $Recall$, and F_1 values obtained by the CKA are usually highest on the datasets of OFFSY, BIEN, Weibo21, and GossipCop. The comparison methods ignore the classification difference among verified news, which restricts their detection performances of fake news. Benefited from the weaken process of the classification disturbance, CKA is reasonable to achieve excellent performances in fake news detection. This confirms the significance of considering the difference of classification knowledge. However, the CKA model's F_1 values shortfall on GossipCop relative to $INNSF$ likely stems from its less effective use of linguistic and stylometric analysis capabilities.

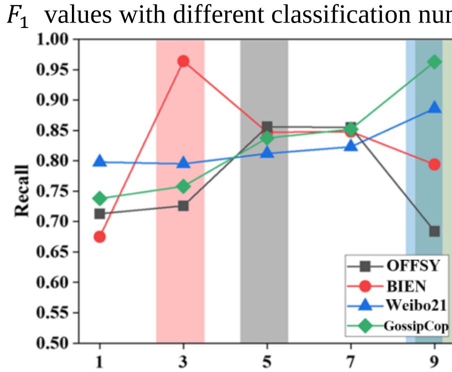
E Parameter Analysis

The classification number is the most important parameter used in our proposed CKA, which is related with the number of classification-specifical models. We discuss the influence of the classification number on the detection performance of fake news in this section. To achieve a controllable setting of the classification number, we flexibly use TwitterNews+ [9] to divide the news items in each training set into different news classifications with different classification values. We use the values of F_1 and $Recall$ as examples to evaluate the detection performance of fake news. As the K value undergoes modifications, Fig. 2 illustrates the alterations in the F_1 and $Recall$ metrics achieved by the CKA across the datasets of OFFSY, BIEN, Weibo21, and GossipCop. The shadows in the background emphasize the best detection performance on each dataset.

As shown in Fig. 2, we can observe that the detection performance of the CKA is strongly correlated with the K value, suggesting the rationality of the proposed CKA. Two important conclusions are summarized as follows. (1) CKA can obtain the best detection performance when the classification number is set as an optimal value. The optimal classification numbers of OFFSY, BIEN, Weibo21, and GossipCop are 5, 3, 9, and 9, respectively. When there are 5, 3, 9 and 9 classification-specifical models in CKA for OFFSY, BIEN, Weibo21, and GossipCop, the corresponding detection performances obtained by CKA are best. (2) When the difference between the setting number and the optimal classification number is larger, the detection performance of the CKA is worse. When the numbers of the classification-specifical models in CKA are closer to the optimal classification numbers, CKA can achieve the higher values of F_1 and $Recall$ on the datasets of OFFSY, BIEN, Weibo21 and GossipCop.



(a) The F_1 values with different classification numbers.

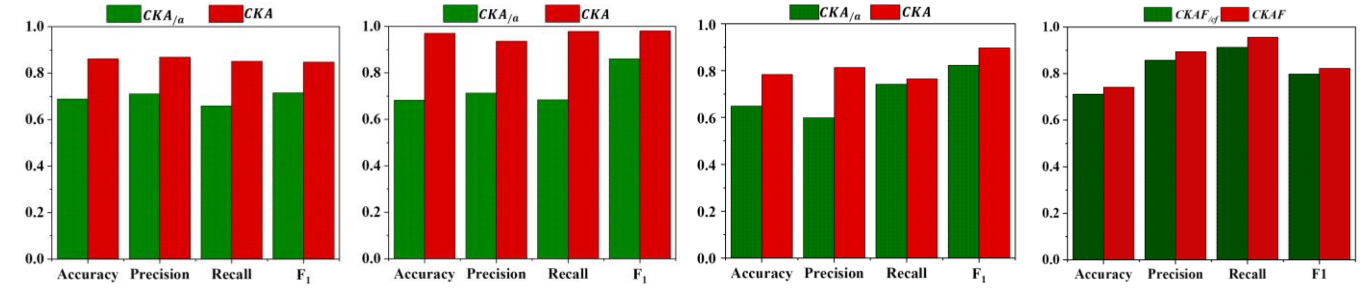


(b) The $Recall$ values with different classification numbers.

Fig. 2. Performances of CKA with different classification numbers on different datasets.

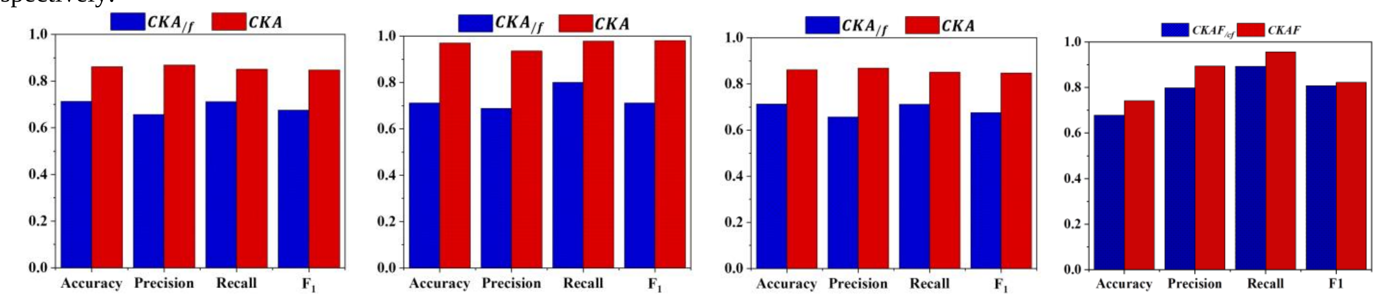
F Importance of Ensemble Coordinator

To demonstrate the importance of the ensemble coordinator, we design a variant of the proposed CKA for comparison. To enable our proposed CKA to exploit multi-classification knowledges during the training stage, we require the ensemble coordinator to minimize the detection loss and maximize the correlation loss. However, even without this maximization process to reduce the correlation among classification-specifical models, our framework can still be trained to generate different classification-specifical models to distinguish between fake and true news. Consequently, a derivative version of the initially proposed architecture has been developed, denoted as $CKA_{/a}$. The distinction between CKA and $CKA_{/a}$ lies in the latter's omission of the correlation loss maximization procedure within the ensemble coordinator. Averaged over 30 independent runs, the performance comparisons between the CKA and $CKA_{/a}$ are shown in Fig. 3.



(a) OFFSY (b) BIEN (c) Weibo21 (d) GossipCop

Fig. 3. Performance comparisons between CKA and $CKA_{/a}$ on the datasets of OFFSY, BIEN, Weibo21 and GossipCop, respectively.



(a) OFFSY (b) BIEN (c) Weibo21 (d) GossipCop

Fig. 4. Performance comparisons between CKA and $CKA_{/f}$ on the datasets of OFFSY, BIEN, Weibo21 and GossipCop, respectively.

As shown in Fig. 3, compared with $CKA_{/a}$, the CKA achieves significant performance improvements in fake news detection, and its improvements of the values of $Accuracy$, $Precision$, $Recall$, and F_1 are obvious. Without the maximization process of correlation loss in the ensemble coordinator, the $CKA_{/a}$ only minimizes the detection loss during the training stage, where each classification-specifical model learns general knowledge by capturing both classification-specifical and sharing features. By contrast, with the supplemental help of the maximization process of the correlation loss, the CKA enables each classification-specifical model to learn discriminative classification-specifical knowledge and keep the classification independence among trained classification-specifical models. It requires the trained classification-specifical models to reduce their classification correlations by removing the sharing features among the news on different news classifications, which enable CKA to resist the potential classification disturbance during the training stage. The comparison results prove the importance of the ensemble coordinator in the CKA, and resisting the classification disturbance during the training stage is beneficial to improving the detection performances.

G Importance of Truth Predictor

In this section, we analyze the importance of the truth predictor. To exploit multi-classification knowledges during the prediction stage, our proposed CKA designs a unified credibility mechanism to fuse the detection conclusions from different classification-specifical models for each unverified news in a truth discovery manner. Without this credibility mechanism in the truth predictor, our framework can still be trained to distinguish between fake and true news. Thus, we design a variant of the proposed method CKA, which is termed $CKA_{/f}$. The $CKA_{/f}$ only removes the credibility mechanism in the truth predictor by assigning each classification-specifical model a fixed and equivalent reliability. In this experiment, the evaluated values of $Accuracy$, $Precision$, $Recall$, and F_1 are averaged over 30 independent runs for each dataset. The performance comparisons between the CKA and $CKA_{/f}$ are shown in Fig. 4.

Fig. 4 illustrates that the CKA achieves better detection performances of fake news on the datasets of OFFSY, BIEN, Weibo21, and Twitter, significantly outperforming the $CKA_{/f}$. This strongly indicates that resisting the classification disturbance during the prediction stage effectively improves the detection performances of fake news. After the ensemble coordinator learns classification-specifical knowledge to generate different classification-specifical models, the truth predictor in the CKA fuses their classification-specifical detection conclusions during the prediction stage though a resilient fusion mechanism. On one hand, the truth predictor evaluates the reliability of each classification-specifical model to enhance the importance of the reliable classification-specifical models. Each classification-specifical model invests its reliability into its detection labels, and the detection label collects its vote score from the invested reliabilities by the classification-specifical models. On the other hand, the truth predictor analyzes the classification correlations among news to enhance the importance of the classification-specifical models on their familiar news classifications. We calculate the classification correlation between the unverified news and the classification-specifical models to adaptively calculate the fake likelihood of each verified news. However, without the unified credibility mechanism in the truth predictor, the $CKA_{/f}$ loses the ability to differentiate between the reliabilities of different classification-specifical models, which severely degrades its detection performances. Therefore, the design of the truth predictor to aggregate classification knowledge during the prediction stage is important and beneficial in the CKA.

VI. Conclusions and Future Work

In this research, we exploit multi-classification knowledges for fake news detection. We propose a multi-classification division-aggregation framework for fake news detection, consisting of three main components, i.e. the news characterizer, the ensemble coordinator, and the truth predictor. The news characterizer extracts news features and obtains news classifications. Cooperating with the news characterizer, the ensemble coordinator generates different classification-specifical models for different news classifications to learn specifical classification knowledge. Finally, the truth predictor aggregates the classification knowledges from different classification-specifical models to predict the fake likelihoods of unverified news. Extensive experimental investigation shows that the exploration of multi-classification knowledges is beneficial in fake news detection and our proposed framework outperforms state-of-the-art comparison methods in fake news detection.

In the future, we plan to explore the fusion of specific features and sharing features among different classifications to improve the detection performance of fake news. Some existing work can be referred for consultation [38]. We aim to utilize advanced natural language processing algorithms to extract key information and create more uniform representations, ensuring that the varying lengths of news do not lead to sparsity in the representations. Additionally, we will consider incorporating more evaluation metrics to enhance the quality of our experiments [39]. Further, we will use a pre-trained CNN version in the news characterizer and compare its performance to our current approach.

References

[1] C. Li, H. Peng, J. Li, L. Sun, L. Lyu, L. Wang, P. Yu, and L. He, "Joint stance and rumor detection in hierarchical heterogeneous graph," *IEEE transactions on neural networks and learning systems*, vol. 33, no. 6, pp. 2530-2542, 2022.

[2] Q. Liao, H. Chai, H. Han, X. Zhang, X. Wang, W. Xia, and Y. Ding, ''An integrated multi-task model for fake news detection,'' *IEEE Transactions on Knowledge and Data Engineering*, pp. 1-12, 2021.

[3] K. Ali and K. Zain-Ul-Abdin, ''Post-truth propaganda: heuristic processing of political fake news on Facebook during the 2016 U.S. presidential election,'' *Journal of Applied Communication Research*, pp. 1-20, 2020.

[4] X. Chen, F. Zhou, F. Zhang, and M. Bonsangue, "Catch me if you can: A participant-level rumor discovery framework via fine-grained user representation learning," *Information Processing and Management*, vol. 58, no. 5, pp. 102678, 2021.

[5] Y. Wang, W. Yang, F. Ma, J. Xu, B. Zhong, Q. Deng, and J. Gao, ''Weak supervision for fake news detection via reinforcement learning," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 1, pp. 516-523, 2020.

[6] Q. Nan, J. Cao, Y. Zhu, Y. Wang, and J. Li, "MDFEND: Multi-domain fake news detection," in *Proceedings of the ACM International Conference on Information and Knowledge Management*, pp. 3343-3347, 2021.

[7] F. Tan, C. Hu, Y. Hu, K. Yen, Z. Wei, A. Pappu, S. Park, and K. Li, "MGEL: Multigrained representation analysis and ensemble learning for text moderation," *IEEE transactions on neural networks and learning systems*, pp. 1-10, 2022.

[8] J. Yang and W. Tay., "An unsupervised Bayesian neural network for truth discovery in social networks," *IEEE Transactions on Knowledge and Data Engineering*, vol. 34, no. 11, pp. 5182-5195, 2022.

[9] M. Hasan, M. A. Orgun, and R. Schwitter, "Real-time event discovery from the Twitter data stream using the TwitterPost+ Framework," *Information Processing and Management*, vol. 56, no. 3, pp. 1146-1165, 2018.

[10] X. Wang, T. Hu, and L. Tang, "A multiobjective evolutionary nonlinear ensemble learning with evolutionary feature selection for silicon prediction in blast furnace," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 5, pp. 2080-2093, 2022.

[11] D. Zhang, D. Wang, N. Vance, Y. Zhang, and S. Mike, " On Scalable and Robust Truth Discovery in Big Data Social Media Sensing Applications," *IEEE Transactions on Big Data*, vol. 5, vol. 2, pp. 195-208, 2019.

[12] H. Shao, D. Sun, S. Yao, L. Su, Z. Wang, D. Liu, S. Liu, L. Kaplan, and T. Abdelzaher, "Truth discovery with multi-modal data in social sensing," *IEEE Transactions on Computers*, vol. 70, no. 9, pp. 1325-1337, 2020.

[13] H. Xiao, J. Gao, Q. Li, F. Ma, L. Su, Y. Feng, and A. Zhang, "Towards confidence interval estimation in truth discovery," *IEEE Transactions on Knowledge and Data Engineering*, vol. 31, no. 3, pp. 575-588, 2018.

[14] Y. Li, J. Gao, C. Meng, Q. Li, L. Su, B. Zhao, W. Fan, and J. Han, "A survey on truth discovery," *ACM Sigkdd Explorations Postletter*, vol. 17, no. 2, pp. 1-16, 2016.

[15] D. M. J. Lazer, M. A. Baum, Y. Benkler, A. J. Berinsky, K. M. Greenhill, F. Menczer, M. J. Metzger, B. Nyhan, G. Pennycook, D. Rothschild, M. Schudson, S. A. Sloman, C. R. Sunstein, E. A. Thorson, D. J. Watts, and J. L. Zittrain, "The science of fake news," *Science*, vol. 359, no. 6380, pp. 1094-1096, 2018.

[16] A. Kucharski, "Post-truth: Study epidemiology of fake news," *Nature*, vol. 540, no. 7634, pp. 525-525, 2016.

[17] H. Ahmed, I. Traore, and S. Saad, "Detection of online fake news using n-gram analysis and machine learning techniques," *in Proceedings of the International Conference on Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments*, pp. 127-138, 2017.

[18] G. Gravanis, A. Vakali, K. Diamantaras, and P. Karadais, "Behind the cues: A benchmarking study for fake news detection," *Model Systems with Applications*, vol. 128, pp. 201-213, 2019.

[19] N. K. Conroy, V. L. Rubin, and Y. Chen, "Automatic deception detection: Methods for finding fake news," *in Proceedings of the Association for Information Science and Technology*, vol. 52, no. 1, pp. 1-4, 2015.

[20] G. Nguyen, S. Dlugolinsky, M. Bobak, V. Tran, Á. L. García, I. Heredia, P. Malík, and L. Hluchý, "Machine learning and deep learning frameworks and libraries for large-scale data mining: a survey," *Artificial Intelligence Review*, vol. 52, no. 1, pp.77-124, 2019.

[21] Y. Liu and Y. Wu, "FNED: A deep network for fake news early detection on social media," ACM Transactions on Information Systems, vol. 38, no. 3, pp.1-33, 2020.

[22] J. Ma, W. Gao, P. Mitra, S. Kwon, B. J. Jansen, K. F. Wong, and M. Cha, "Detecting rumors from microblogs with recurrent neural networks," in *Proceedings of the International Joint Conference on Artificial Intelligence*, pp. 3818-3824, 2016.

[23] K. Shu, D. Susan, H. Ahmed, and L. Huan, "Detecting fake news with weak social supervision," *IEEE Intelligent Systems*, vol. 36, no. 4, pp. 96-103, 2020.

[24] C. Song, Y. Teng, Y, Zhu, S. Wei, and B. Wu, "Dynamic graph neural network for fake news detection," *Neurocomputing*, vol. 505, pp. 362-374, 2022.

[25] M. Davoudi, M. Moosavi, and M. Sadreddini, "A hybrid deep model for fake news detection using propagation tree and stance network," *Expert Systems with Applications*, vol. 198, 116635-116635, 2022.

[26] P. Meel and D. Vishwakarma, "A temporal ensembling based semi-supervised ConvNet for the detection of fake news articles," *Expert Systems with Applications*, vol. 177, pp. 115002, 2021.

[27] L. Wu, Y. Rao, C. Zhang, Y. Zhao, and A. Nazir, "Category-Controlled Encoder-Decoder for Fake News Detection," *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 2, pp. 1242-1257, 2023.

[28] J. Li, H. Xie, L. Yu, X. Gao, Y. Zhang, "Discriminative feature mining based on frequency information and metric learning for face forgery detection," *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 12, pp. 12167-12180, 2023.

[28] Z. Shang, H. Xie, L. Yu, Z. Zha, Y. Zhang, "Constructing Spatio-Temporal Graphs for Face Forgery Detection," *ACM Transactions on the Web*, vol. 17, no. 3, pp. 1-25, 2023.

[30] H. C. Shin, H. R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. Mollura, and R. M. Summers, "Deep convolutional neural networks for computeraided discovery: CNN architectures, dataset characteristics and transfer

learning," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1285-1298, 2016.

[31] D. Zhang, H. Xu, Z. Su, and Y. Xu, "Chinese comments sentiment classification based on word2vec and SVMperf," *Model Systems with Applications*, vol. 42, no. 4, pp. 1857-1863, 2015.

[32] X. Yin, J. Han, and P. S. Yu, "Truth discovery with multiple conflicting information providers on the web," *IEEE Transactions on Knowledge and Data Engineering*, vol. 20, no. 6, pp. 796-808, 2008.

[33] Y. Wang, F. Ma, Z. Jin, Y. Yuan, G. Xun, K. Jha, L. Su, and J. Gao, "EANN: Event adversarial neural networks for multi-modal fake news detection," *in Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 849-857, 2018.

[34] Biendata Competitions,https://www.biendata.xyz/competition/fakepost /?source=bdzhjg1.

[35] K. Shu, D. Mahudeswaran, S. Wang, D. Lee, and H. Liu. "FakeNewsNet: a data repository with news content, social context, and spatiotemporal information for studying fake news on social media," *Big Data*, vol. 8, no. 3, pp. 171-188, 2020.

[36] Z. Wang and J. Du, "Joint architecture and knowledge distillation in CNN for Chinese text recognition," *Pattern Recognition*, vol. 111, pp. 107722, 2021.

[37] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," *Computer Science*, 2014.

[38] Z. Gao, C. Sun, Z. Cheng, W. Guan, A. Liu, and M. Wang, "TBNet: A two-stream boundary-aware network for generic image manipulation localization," *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 7, pp. 7541-7556, 2022.

[39] Y. Zhu, Q. Sheng, J. Cao, Q. Nan, K. Shu, M. Wu, J. Wang, and F. Zhuang, "Memory-guided multi-view multi-domain fake news detection," *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 7, pp. 7178-7191, 2023.