An Entity Ontology-Based Knowledge Graph Embedding Approach to News Credibility Assessment

Qi Liu[1], *Senior Member, IEEE*, Yuanyuan Jin[1],Xuefei Cao,Xiaodong Liu *Senior Member, IEEE*, Xiaokang Zhou[*], *Member, IEEE*, Yonghong Zhang, Xiaolong Xu and Lianyong Qi

***Abstract*—Fake news is a prevalent issue in modern society, leading to misinformation and societal harm. News credibility assessment is a crucial approach for evaluating the accuracy and authenticity of news. It plays a significant role in enhancing public awareness and understanding of news, while also effectively mitigating the dissemination of fake news. However, news credibility assessment meets challenges when processing large-scale and constantly growing data, due to insufficient and unreliable labels and standards, and diversity and semantic ambiguity of news contents. Recently, machine learning models have been well developed to address these issues, but suffer from limited effectiveness. A unified framework is also required for them to represent various entities and relationships involved in news stories. This paper proposes an Entity Ontology-Based Knowledge Graph Network (EKNet) to leverage knowledge graphs and entity frameworks for news credibility assessment. The model utilizes the information from knowledge graphs by combining entities and relationships from news and knowledge graphs. Experimental results show that the EKNet has advantages in evaluating news credibility over existing methods. Specifically, compared to several strong baselines, the model demonstrates a significant performance improvement in scores across various tasks. Which indicates that using the EKNet to address the challenges in news credibility assessment is highly effective and can conduct better performance for the problem of fake news in the social media environment.**

***Index Terms*—Fake news detection, News credibility assessment, Knowledge Enhancement, Knowledge graph.**

This work has received funding from National Key Research and Development Program of China under Grant No.2021YFE0116900, National Natural Science Foundation of China (No. 42275157 and 41975142), Major Program of the National Social Science Fund of China (Grant No. 17ZDA092), and basic Research Programs (Natural Science Foundation) of Jiangsu Province (BK20191398).

Q. Liu is with the School of Software, Nanjing University of Information Science and Technology, Nanjing 210044, China (e-mail: qi.liu@nuist.edu.cn).

Y. Jin is with the School of Software, Nanjing University of Information Science and Technology, Nanjing 210044, China (e-mail: echo0301@qq.com).

X. Cao is with the School of Cyber and Information Security, Xidian University, Xi'an, 710071, China (e-mail: xfcao@xidian.edu.cn).

X. Liu is with the School of Computing, Edinburgh Napier University, Edinburgh EH10 5DT, U.K. (e-mail: x.liu@napier.ac.uk).

X. Zhou is with the Faculty of Data Science, Shiga University, Hikone 522-8522, Japan, and also with the RIKEN Center for Advanced Intelligence Project, RIKEN, Tokyo 103-0027, Japan. (e-mail: zhou@biwako.shiga-u.ac.jp).

Y. Zhang is with the School of Automation, Nanjing University of Information Science & Technology, No. 219, Ningliu Road, Nanjing, 210044, China.

X. Xu is with the School of Software, Nanjing University of Information Science and Technology, Nanjing 210044, China (e-mail: xlxu@ieee.org).

L. Qi is with the School of Information Science and Engineering, Qufu Normal University, Rizhao, Shandong China (e-mail: lianyongqi@gmail.com).

[1]Both authors are the first author due to equal contribution to this paper.

[*]The correspondence author of this paper.

I. Introduction

The rise of digital media has made news and information more accessible than ever. While this is undoubtedly a positive development for journalism, the sheer volume of news articles online can make it difficult for readers to determine which sources are trustworthy and which are not. Fake news has become a serious problem in today's social media environment. Fake news refers to false or misleading information based on misleading headlines and content that is widely disseminated and causes public panic and misunderstanding. It not only damages the credibility of the media but also the whole country and society. It is increasingly difficult for readers to determine which sources provide high-quality, trustworthy news. In this context, fake news detection and assessing news credibility becomes particularly important. However, these tasks face multiple challenges.

Given this backdrop, the task of detecting fake news and evaluating news credibility assumes paramount importance. However, these endeavors encounter several challenges. One such challenge is the presence of semantic ambiguity, adding complexity to the identification of fake news and the assessment of news credibility.

To address these challenges, researchers have proposed various workarounds. One solution is a deep learning-based approach[2], [3], including convolutional neural networks (CNN) and recurrent neural networks (RNN), for processing text to improve detection of fake news.

In fake news, fictitious entities and false relationships are often involved, which are difficult to be captured by traditional models[1]. If only traditional models such as convolutional neural network (CNN) and recurrent neural network (RNN) are used, the model will only be able to learn surface features from the text itself, but cannot capture the relationship and semantic information between entities. This leads to the fact that the detection accuracy of fake news may not be high. It is difficult for traditional models to deal with entities with ambiguity or complex structures accurately. For example, many entities have multiple meanings or may have different meanings in different contexts, traditional models may treat them as different entities, thus affecting the accuracy of the model. Some entities described in the news may have complex structures, such as hierarchical structures, cross-references, etc., which will also increase the processing difficulty of the model. In addition, most fake news discrimination models will only tell you whether the news is fake, but will not output the credibility and evaluation basis of the news.

To address the above limitations, this paper proposes a new fake news credibility assessment method EKNet(Entity Ontology-Based Knowledge Graph Network), which utilizes knowledge graph and entity ontology framework to assess the credibility of news. Specifically, the model introduces external knowledge as the basis for fake news evaluation. This paper also proposes an entity ontology-based knowledge base storage and management framework, and points out the evaluation basis of news through this framework. The model pointed out inaccuracies in the news. In addition, since the dataset contains Chinese data, a Chinese word segmentation method is proposed. In summary, the contributions of this paper are as following:

1) An Entity Ontology Framework: The Entity Ontology Framework is proposed for storing and representing entity ontologies. The Entity Ontology Framework comprises two parts: a representation model and an entity ontology repository.
2)EKNet: A novel approach, EKNet, that combines a knowledge graph encoder with a text encoder to detect fake news and generate comments on news articles.
3)Integration of an entity ontology framework into the EKNet architecture, allowing for more nuance analysis of relationships between entities in news articles.

In this paper, recent contributions and near-future trends on news credibility assessment are discussed. Section 2 discusses related works. Section 3 and section 4 describe our approach to the evaluation of news credibility. The corresponding experiment and evaluation are then presented in Section 5. In Section 6, a conclusion and future work are discussed.

II. Related Work

A. News Credibility Assessment

Due to the widespread concern of deep learning [32], [33], there has been a growing interest in developing methods for news credibility assessment. The majority of these methods utilize machine learning algorithms to identify and evaluate various features that can be used to distinguish between credible and non-credible news. Several studies have focused on using textual features, such as linguistic style, sentiment analysis, and content-based features, to classify news articles as credible or not [20], [21], [4], [8]. Kai et al.[7] proposed a hierarchical model to detect the confidence of sub-events in news reports to judge the quality of the whole news. Gupta et al. [9] proposed a method to analyze the credibility of news based on the response time, validity and availability of news. Conroy et al. [10] put forward a kind of the user's behavior data combined with content of fake news detection methods. ElAzab et al. [11] put forward a kind of fake news on the modeling of an object recognition based on the propagation behavior model, wandering in the spread of fake news trajectory tracking, and through the graph model and evolution model for specific fake news in further. In this method, the recognition of fake news is regarded as a long text classification problem. The recurrent neural network RNN, LSTM[12], or GRU receives each sentence from the news and processes it. The recurrent neural network's hidden layer vector is then used to represent the news sentence embedding, and the hidden layer information is then input into the classifier to get the classification result. Ruchansky et al. [5] modeled the three characteristics of news information, news publishers and users' reactions to it, and modeled fake news. Ma, Jing et al.[6] believed that the authenticity of news was closely related to the position expressed in response posts. They proposed to combine the position classification and fake news detection into a multi-task model, and applied multi-task to fake news detection for the first time.

Other approaches incorporate social media data as additional features for credibility assessment. For instance, Liao et al.[30]proposed a method that uses features extracted from the Twitter network, including user interactions and tweet content, to classify news articles. Similarly, Gupta et al. [9] used social media data to predict the veracity of news articles by analyzing the propagation patterns of tweets.

Despite the progress made in news credibility assessment, there are still several challenges that need to be addressed. One of the main challenges is the lack of labeled data, as it can be difficult and time-consuming to manually annotate news articles as credible or not. Another challenge is the presence of adversarial examples, where an attacker intentionally modifies the content of a news article to mislead a credibility assessment model [31].

In summary, there have been various approaches proposed for news credibility assessment, ranging from textual features to social media data and knowledge graphs. However, the field still faces several challenges, including the lack of labeled data and the presence of adversarial examples. Addressing these challenges will be crucial for developing more accurate and reliable methods for news credibility assessment.

B. Knowledge Graph Enhancement

Researchers have explored different domains and approaches to leverage knowledge graphs and improve the performance of deep learning models.

Jiang et al. [29] introduce "knowledgeable prompt learning (KPL)" for fake news detection, leveraging and external knowledge to outperform traditional fine-tuning of pretrained language models. Peters et al. [22] propose a method to enhance contextual word representations with structured knowledge from multiple knowledge bases. Similarly, Hao et al. [23] introduce Sentiment Knowledge Enhanced Pre-training for sentiment analysis tasks, leading to improved performance on sentiment analysis tasks.

Shreyansh et al. [24] focus on knowledge graph semantic enhancement for improving artificial intelligence. They discuss the integration of background knowledge represented in knowledge graphs with machine learning algorithms, enhancing accuracy and explaination in recommendation and community detection tasks. Dehai et al. [25] propose a Knowledge Graph Based Synthetic Corpus Generation approach to enhance language model pre-training.

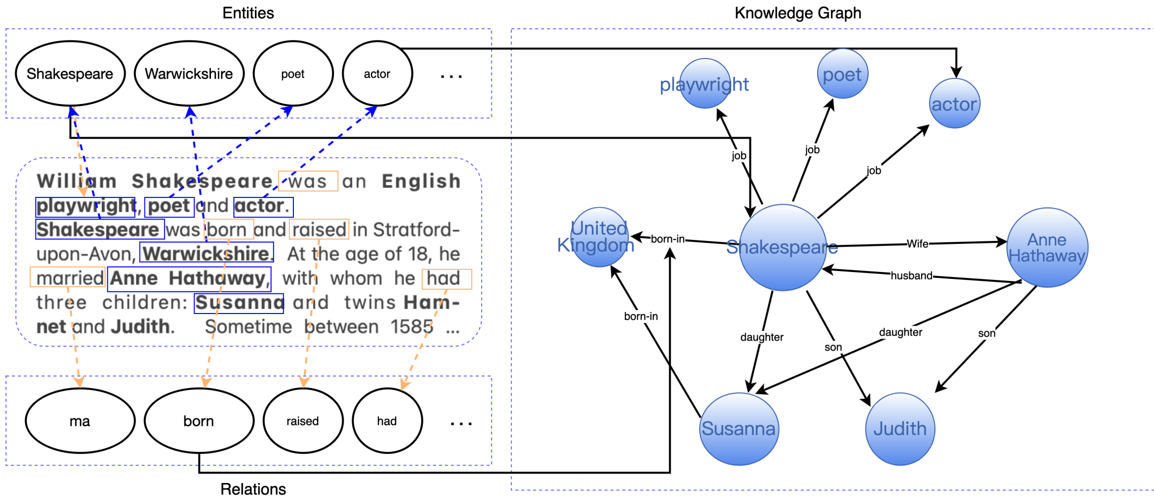

Fig. 1. An example of a knowledge graph containing entities and relationships in the news.

Furthermore, knowledge graph-enhanced neural collaborative recommendation is introduced by Lei et al. [26] to address the sparsity issue in neural collaborative filtering methods. In a different context, Linmei et al. [27] explore text-graph enhanced knowledge graph representation learning, utilizing heterogeneous entity-text graphs to enhance knowledge graph embeddings with global word co-occurrence information.In the medical domain, Zhixue et al. [28] design a medical question answering system based on a knowledge graph.

In summary, researchers have explored various approaches to enhance artificial intelligence tasks using knowledge graph integration. These methods range from sentiment analysis, recommendation systems, and medical question answering to language model pre-training.

III. Entity Ontology Framework

A. Overview of Entity Ontology Framework

In order to address the challenges posed by the increasing complexity of news articles and the demand for efficient, accurate, and news credibility assessment, as well as to better manage the structure and information of all entity ontologies. This paper proposes a method for storing and representing entity ontologies, referred to as the Entity Ontology Framework. The Entity Ontology Framework consists of two components: a representation model and an entity ontology repository. This model defines all entity ontologies in news texts and establishes a hierarchical structure to categorize entity parameters. These predefined parameters can then be maintained and shared via the repository. The Entity Ontology Framework contributes to consistency and standardization among news articles. The establishment of a hierarchical structure and the use of predefined parameters ensure consistent categorization, which is crucial for maintaining uniform scoring and evaluation practices across various news topics and sources.

The construction of the Entity Ontology for News Articles involves several steps. As is shown in figure 1, all the entities mentioned in the news text are identified and extracted. Then, a hierarchical structure is established to categorize entity parameters, and predefined parameters are maintained and shared via the repository. This approach helps to improve the fineness of entity messages and reduce the complexity of entity ontology representation. EKNet(Entity Ontology-Based Knowledge Graph Network)'s entity ontology framework can improve the accuracy and reliability of news credibility evaluation. Specifically, the entity ontology framework provides a structured, comprehensive representation of entities and their relationships, which can be used to infer and verify the credibility of news articles. By utilizing entity ontologies, the EKNet can capture the relationships between different entities mentioned in news articles and their respective attributes. These relationships and attributes can then be used to build a knowledge graph that serves as a structured representation of the information provided in the article.

News content alone cannot accurately judge the truth or falseness of news. The knowledge graph can be judged from many perspectives.

1)Knowledge graph can introduce semantic associations between various entities, which can help to find potential connections and thus improve the accuracy of news identification.

2)The knowledge graph can reasonably extend the external knowledge of the news through type relationships.

3)The knowledge graph can also link to entities in news content to find errors in the news. And the interpretation is strong.

With this knowledge graph, EKNet can perform a variety of credibility evaluations that effectively manage and represent entities and relationships in news articles, leading to more accurate and reliable credibility evaluations.

B. Construction of Entity Ontology for News Articles

The Entity Ontology Framework plays a crucial role in EKNet, a system designed for scoring and evaluating news articles based on their trustworthiness. This section will provide a detailed description of the construction of the entity ontology for news articles and its incorporation into the EKNet architecture.

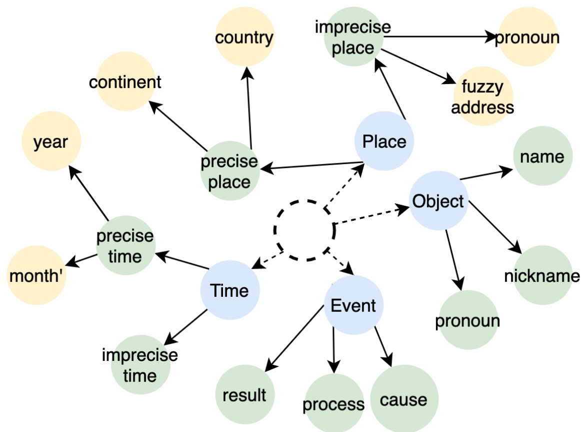

Fig. 2. The predefined entity parameters for managing entities. It describes the structure of four dimensions.

As shown in figure 2, the hierarchical structure consists of four dimensions for storing entity ontologies in the system. In the context of the figure, the blue nodes represent these four dimensions, while the green nodes correspond to secondary structures, and the yellow nodes represent tertiary structures. Each refined parameter has a dotted format index label that represents all the upper dimensions of that dimension. This hierarchical structure is specifically designed to accommodate different levels of entity ontologies, providing a comprehensive framework to represent complex relationships between entities.

In order to comprehensively store and manage all valid entity-related information within the system, the Entity Ontology Base (EoB) is meticulously designed and comprises two integral data structures: EoBSchema and EoBData. EoBSchema is specifically designed to store hierarchical entity dimension information, which encompasses the hierarchical representation of entities within the system. This structure enables the organization and classification of entities based on their attributes and relationships. On the other hand, EoBData serves as a fundamental data structure responsible for storing all valid entities. It effectively functions as a repository for a collection of entity triples, denoted as (e_i, p_j, v_k). Here, e_i represents the i^{th} entity found within the news text, p_j signifies the relationship that links the i^{th} entity with other related entities, and v_k corresponds to the entity value associated with the aforementioned relationship p_j. Each triplet within EoBData essentially signifies an edge in the knowledge graph, collectively forming the entire entity relationship graph, which is pivotal for understanding the intricate web of relationships between entities. The overall size of an EoB instance is contingent upon the dimensional intricacies presented within the news text and the quantity of entities contained within the dataset, ensuring scalability and adaptability to various knowledge domains.

The Entity Ontology Framework is incorporated into EKNet architecture, allowing for more accurate and effective scoring and evaluation of news articles. The framework enables the EKNet to capture the semantic relationships between entities in the news text and helps to improve the overall performance of the system.

In summary, the Entity Ontology Framework provides a comprehensive and efficient approach for managing entity ontologies in news articles. By incorporating this framework into EKNet architecture, the system can better capture and analyze the underlying structure and information of news articles, resulting in more accurate and trustworthy news evaluation.

The news text is long, the model first divides the news sentences. As shown in figure 3, word segmentation operation is also needed, because it is a Chinese dataset. The method of word segmentation will be introduced in the following section. After word segmentation, model use tf-idf to calculate the weight of all words in every sentence, and then calculates the weighted mean of each sentence's word embedding, and the result is regarded as sentence embedding. Then, the weighted average of sentence embedding is calculated to obtain the news embedding.

IV. EKNet: Incorporating Entity Ontology Framework for News Credibility Assessment

This chapter introduces the EKNet(Entity Ontology-Based Knowledge Graph Network), a model that leverages a knowledge graph to improve the accuracy of news verification and generate concise comments for fake news. Building on the entity ontology framework presented in Chapter 3, the EKNet integrates external knowledge from the knowledge graph to augment the representation of entities in news articles.

The first section of this chapter describes how EKNet utilizes the entity ontology framework to extract entity features and establish entity relations from news text. The EKNet then employs a knowledge graph to enrich the entity representation with additional information and infer relationships between entities. By integrating entity features and relations from the ontology framework and knowledge graph, the EKNet constructs a comprehensive representation of news articles that facilitates accurate verification.

The second section focus on how EKNet generates comments for fake news. The EKNet utilizes the same entity ontology framework and knowledge graph to identify key entities and their relations in a news article. From this representation, the EKNet generates concise comments that highlight the problematic aspects of the news.

A. *News Evaluation using EKNet*

The EKNet architecture comprises three main components: a text encoder, a knowledge graph encoder, and a decision maker.

The text encoder is tasked with converting news articles into vector representations. Given that news texts are often lengthy and intricate, our model initially segments the news article into sentences. While it's worth noting that one of the datasets employed in this study is in Chinese, requiring a preliminary word segmentation step, our approach isn't limited to Chinese text. After segmenting the words, the tf-idf method is employed to compute the weight of each word

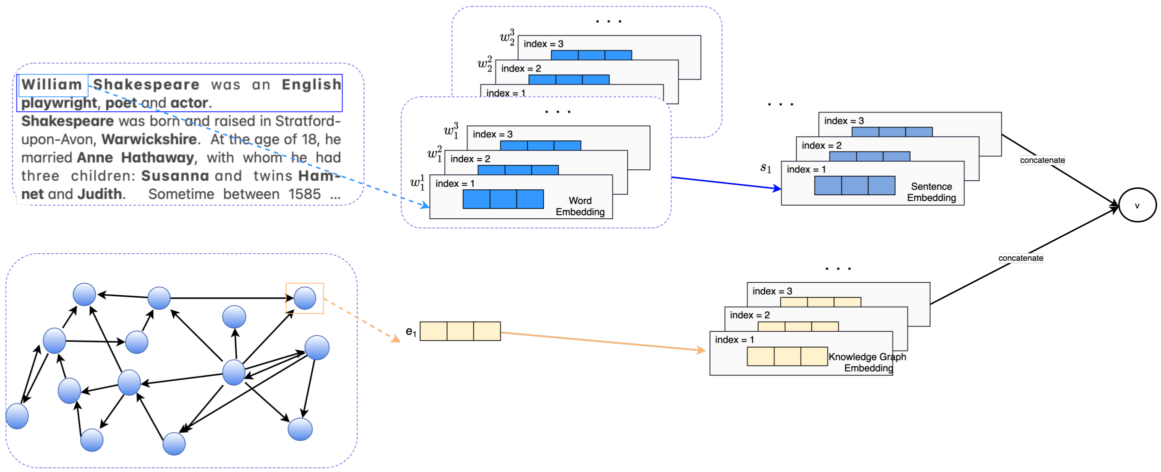

Fig. 3. Alignment and fusion of knowledge graph entity vectors with entity vectors from news articles.

in every sentence. Subsequently, the weighted average of word embeddings within each sentence is calculated. This process results in a vector representing the sentence. Finally, the weighted average of all sentence embeddings is computed to obtain the overall embedding of the news article.

The knowledge graph encoder, on the other hand, encodes the entity ontology and its relationship with Baidupedia into a knowledge graph embedding. This process involves linking each entity in the news text with Baidupedia to obtain its entity relation triplet, from which entity embedding can be calculated.

The decision maker takes both the news article vector and the knowledge graph embedding as input and outputs a binary classification result, indicating whether the news article is real or fake. By incorporating the Entity Ontology Framework, EKNet can capture the complex relationships between entities in news articles and make informed decisions about their credibility.

B. Comment Generation using EKNet

To further enhance the interpretability of EKNet, the EKNet introduce a comment generator component that produces short comments about the news article's credibility. The comment generator takes the news article vector and the knowledge graph embedding as input and generates a short comment as output. Unlike traditional text generation models, the comment generator in the EKNet is designed to generate comments that are informative and relevant to the credibility of the news article. Specifically, the generated comments aim to highlight any dubious or false information that may be present in the news article.

To achieve this, EKNet modify the architecture to include the comment generator component. The comment generator is based on a Encoder-Decoder framework with a BiLSTM encoder and an LSTM decoder. The input to the comment generator is the news article vector and the knowledge graph embedding. The attention mechanism in the Encoder-Decoder model allows the model to focus on relevant parts of the input sequence, which addresses the issue of generating high-quality comments for long articles. The encoder takes in the news article vector and the knowledge graph embedding, and produces a sequence of hidden states that capture the relevant information about the news article. The text's tokens are fed into the encoder one by one, which results in a string of hidden states for the encoder h_i. The decoder receives the word embedding of the preceding word in each step t (during training, t represents the preceding word of the reference digest; during testing, t represents the preceding word emitted by the decoder) and is in the state of the decoder d_t.

This is how the attention distribution a^t is calculated:

$$e^t = p^t tanh(W_h h_i + W_s d_t + b_{attn}) \tag{1}$$

$$a^t = softmax\left(e^t\right) \tag{2}$$

Where p, W_h, W_s, and b_{attn} are variables that can be learned. The context vector h_t^* is created by weighting the hidden states by b_{attn} attention distribution, of the encoder:

$$h_t^* = \sum_i a_i^t h_i \tag{3}$$

The word distribution P_{vcb} is obtained by the decoder, and the context vector is connected to the encoder state d_t:

$$P_{vcb} = softmax(V'(P[d_t, h_t^*] + b) + b') \tag{4}$$

P, P', b and b' are parameters learned from the model. V' is a learned parameter mapping decoder states to the vocabulary feature space. The final distribution from which words w can be predicted is P_{vcb} , the probability distribution of every word in the vocabulary:

$$P\left(w\right) = P_{vcb}\left(w\right) \tag{5}$$

The negative log probability of the target word w_t^* for timestep t during training is the loss:

$$loss_t = -logP\left(w_t^*\right) \tag{6}$$

The total loss of the whole generated sequence is:

$$total = \frac{1}{N}\sum_{t=0}^{N} loss_t \tag{7}$$

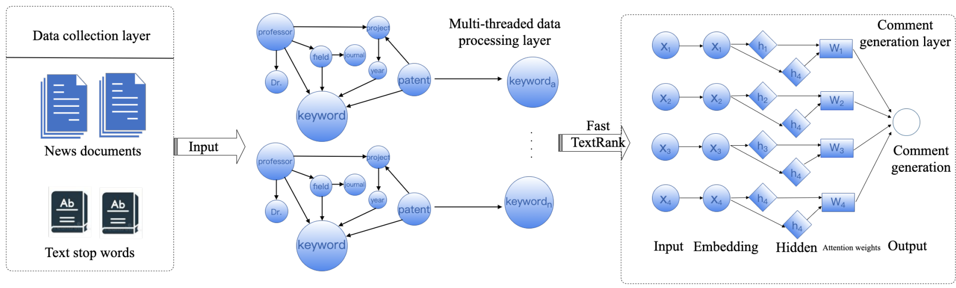

Fig. 4. Data preprocessing using FastTextRank algorithm.

The decoder then generates the output sequence one token at a time, conditioned on the hidden states from the encoder.

C. Data Processing

1) Data Enhancement: The process of data preprocessing is shown in figure 4. The FastTextRank algorithm is used in data preprocessing to automatically extract a number of meaningful words or phrases from the given text of the dataset and write them to a text file. The FastTextRank algorithm is used to rank subsequent keywords using relationships between local words (co-occurrence windows), extracted directly from the text itself.

The data enhancement processes are shown in Algorithm 1. The algorithm takes as input the $Text$ which represents the text to be processed, $keywords$ indicating the presence of respective keywords, and $Synonyms$ denoting a list of synonyms. For each index in $Text$, the algorithm checks for the presence of $keywords$. If they exist, it replaces these keywords with their corresponding $Synonyms$. Subsequently, the algorithm transforms the text at the current index, $Text_{index}$, into a new word, $wordNew$, using a coverage mechanism. If $wordNew$ is not already in the generated vocabulary $Generatingwords$, it is added to this list, and the current sentence is appended to $Text_{sentences}$. If $wordNew$ is already in $Generatingwords$, it is not added. The algorithm repeats these steps until it traverses all indices in $Text$. Upon completion of the algorithm's execution, it yields the $Generatingwords$ and $Textsentences$, signifying lists of newly generated words and the corresponding sentences, respectively. These lists encapsulate the outcomes of the algorithm's processing on the provided $Text$.

Data enhancement can solve the problem of few samples within the NLP domain. This session implements several techniques for data enhancement as follows:

1) Word replacement: Since there is no sophisticated near-synonym dictionary like WordNet available for Chinese, the word vector space of embedding is chosen to find the closest semantic words. By using a pre-trained Chinese word vector on a large amount of data, the closest word to the semantic meaning of each word in that word vector space can be reached, and then the words in the original sample are replaced to obtain a new sample. However, one problem is that if the core vocabulary in the sample is replaced, it may result in the loss of core semantics. There are two solutions to this problem: the first one is to rank the words in the word list by tf-idf weights, and then replace the words with the lowest ranking. The second one is to first mine the topic words in the sample by unsupervised way, and then replace only the words that do not belong to the topic words.

Algorithm 1 Data Enhancement

Input: $Text$, $keywords$, $Synonyms$
Output: $Generatingwords$, $Textsentences$

```
For each index in Text do
    If Text_index have keywords
        Set keywords ← Synonyms
    End If
        Set wordNew ← Text_index to Coverage mechanism
    If wordNew not find in Generatingwords
        Generatingwords append to wordNew
        Text append to Textsentences
    Else
        Generatingwords not append to wordNew
    End If
End For
return Generatingwords, Textsentences
```

2) Coverage mechanism: A common issue with sequence-to-sequence models is text repetition, and is more obvious when generated text content has multiple sentences. Using the Coverage mechanism (CVG), the problem of invalid repetition of phrases in the data can be solved.

3) Self-service sample generation: Once a text generation model is trained, the trained model can be used to generate a new source for the reference in the original sample and continue training the model as a new sample.

2) Chinese Word Segmentation: There are many domain-specific nouns in news text. Chinese word segmentation is regards as a single character-based sequence labeling problem in the Chinese word segmentation model utilized in this study, which blends machine learning techniques with dictionaries. The model introduces an external dictionary, which can enhance domain adaptability.

This study defines an array of states called $[D, S, Z, M, l]$, where D represents "Dan", which denotes that a Chinese character is a word. "Shi", the first character of a word, is what S stands for. The character Z is "Zhong", which denotes that it is in the middle of the Chinese word. M is "Mo", which means the last character of a Chinese word. The model scans the sentence to obtain a set of $DSEM$ sequences with the highest probability. Sort the sequence to get the word segmentation result. Chinese sentences are used to putting emphasis on the back. So the model introduced l to represent the position where the word appears, and different positions have different weights.

The Chinese corpus training word vector in the training set, converted into the four labeling forms of $DSZM$, and l is used to mark the position of each word. The corpus is organized into a word dictionary, the word frequency of each word is calculated, the probability table is obtained by combining the position information training. To enhance domain adaptability, external dictionaries are also introduced. The model scans the sentence and queries the dictionary to obtain multiple possible sentence segmentation results, obtains the $DSZM$ sequence with the highest probability through dynamic programming. For words which are out of vocabulary, the model uses $DSZM$ as the hidden state of CRF and combines it with the Viterbi algorithm to obtain the optimal $DSZM$ sequence. Finally, according to the following rules, the word segmentation result is obtained.

1) There is no continuous S or M.
2) There must be S in front of Z, there must be M behind.
3) S and M come in pairs.

V. Results and Evaluation

A. Datasets

The datasets used in this study consist of both Chinese and English texts, reflecting the diverse linguistic context of news content.

1) Chinese Word Segmentation Dataset: The corpus for this study was derived from the second International Chinese Word Segmentation Contest, with significant modifications to enhance its applicability. These modifications included the addition of manually segmented news data, aimed at increasing the dataset's representativeness. The total dataset comprises 21,001 segmented entries.

For the purpose of analysis, this dataset was divided into a test set and a training set. The test set consists of 1,945 Chinese segmented sentences, encompassing 13,148 distinct Chinese words. The average sentence length in this set is 88.80 characters, with an average word length of 2.28 characters. Conversely, the training set includes 19,056 Chinese sentences, featuring 55,729 different Chinese words. Here, the average sentence length is 95.84 characters, and the average word length is 2.68 characters.

2) Baidupedia Dataset: In this paper, datasets of Baidu News and Baidupedia are used to evaluate how well the news credibility assessment in EKNet(Entity Ontology-Based Knowledge Graph Network) model performs. There were 7834 news items in whole dataset, including 7037 news items in the training set and 797 news items in the test set. The average news item in the whole Baidupedia dataset is 2836.53 characters long, the average news item in the training dataset is 2835.43 characters long, and the average news item in the testing dataset is 2846.23 characters long. There are 272.36 triples in each news article on average, 272.42 triples in the training set and 271.80 triples in each news article in the test set.

3) Real or Fake Dataset[35]: This dataset is the public dataset on Kaggle and contains a total of 6060 pieces of news. Real and fake news is 50/50. The dataset contains four labels, id, $title$, $text$ and $label$. In the context of our experiment, with a specific focus on the task of fake news detection, the decision has been made to utilize only the $text$ and $label$ attributes from this dataset.

4) Fake News Detection Dataset[36]: A public dataset published on Kaggle five years ago, contains a total of 3,988 pieces of news. There are 2,121 real news articles and 1,867 fake news articles. This dataset contains four labels, $URLs$, $Headline$, $Body$ and $Label$. Given the research focus on fake news detection, the decision has been made to utilize solely the $Body$ and $Label$ attributes from this dataset.

B. Evaluation Metrics

Automatic assessing of news is evaluated using a set of metrics in Recall-Oriented Understudy for Gisting Evaluation (ROUGE), which has also been applied to evaluate the short text generation. It calculates the F1 value by comparing the automatically generated text with a set of reference texts (usually manually generated), and thus measures the similarity between the automatically generated text and the reference text.

ROUGE has several variants, ROUGE-N (N accepts values 1, 2, 3, 4.), which measures the overlap of n-grams (single tokens) between the reference text and the generated text, and ROUGE-L measures the longest matching word sequence using the Longest Common Subsequence (LCS). ROUGE-W (less commonly used) measures the weighted longest common suffix, and ROUGE-SU (less commonly used) measures co-occurrence statistics based on skip-bigram and unigram.

To evaluate the performance of the model in detecting fake news, the paper employs Precision, Recall, F1 score, and Miss rate as evaluation metrics. In this evaluation, fake news is treated as positive samples while real news is treated as negative samples. Precision measures the proportion of true positives among all positive predictions, while Recall measures the proportion of true positives among all actual positive instances. F1 score is the harmonic mean of Precision and Recall, providing a single measure of the model's overall performance. Finally, the Miss rate (also known as false negative rate) measures the proportion of actual positive instances that are incorrectly classified as negative. These evaluation metrics provide a comprehensive and objective way to assess the effectiveness of the model in detecting fake news.

C. Experimental Settings

To evaluate the performance of the EKNet model on the task of news credibility assessment, this paper conducted experiments on multiple datasets. During the training process, the EKNet model was optimized using the Adam optimizer with a learning rate of 1e-4 and momentum of 0.9. The model employed the cross-entropy loss function to guide the training process. In order to prevent overfitting, an early stopping

TABLE I
FAKE NEWS DETECTION TASK

Method	Baidupedia Dataset			Real or Fake Dataset			Fake News Detection Dataset		
	P(Precision) ↑	**R(Recall)** ↑	**F1** ↑	**P** ↑	**R** ↑	**F1** ↑	**P** ↑	**R** ↑	**F1** ↑
FastText[40]	64.3%	71.3%	73.9%	79.0%	75.9%	76.3%	83.8%	75.1%	78.7%
TextRNN [41]	89.4%	97.3%	93.2%	93.0%	90.2%	91.6%	/	/	/
TextRCNN[41]	90.0%	98.3%	94.0%	90.1%	89.7%	90.3%	86.8%	83.6%	84.7%
Transformer[43]	96.6%	96.8%	96.7%	91.8%	92.4%	92.1%	89.7%	90.1%	89.4%
Dual [42]	96.6%	**99.6%**	98.1%	93.1%	91.7%	93.8%	**94.1%**	91.1%	**92.6%**
EKNet	**99.5%**	99.4%	**99.5%**	**95.5%**	**92.7%**	**94.1%**	90.6%	**91.5%**	91.0%

↑ higher values indicate better performance

strategy was employed, where training was halted and the model with the best validation performance was saved when the validation loss showed no significant decrease over 5 consecutive epochs. Additionally, the EKNet model incorporated common regularization techniques, including dropout and L2 regularization, to enhance the model's robustness and generalization capability.

D. Overall Results

1) Fake News Assessment Experiment: As shown in table 2, the first group of experiments uses RNN [34], [37] as encoder and decoder, which does not introduce attention mechanism and transmits the last hidden layer state obtained from the encoder to the decoder, and the model of the second group of experiments is RNN context [38], which introduces context at the decoder side based on RNN and transmits all states from the encoder to the decoder. The RNN [34], [37] structures in the above 2 groups of experimental models all adopt GRU networks, use Adadelta as the model optimizer, and use cluster search in the decoding process to obtain summary results. The third group of experiments compare proposed method against SummaReranker [39], a state-of-the art method for summarize. The fourth group of experiments used the EKNet model proposed in this paper. The fifth group of experiments added PGN (Pointer generator network) to the EKNet, and the sixth group of experiments added PGN and CVG (Coverage mechanism) to the EKNet. PGN helps to copy words from the source text by pointers, which improves the OOV (Out-of vocabulary) word accuracy and processing while retaining the ability to generate new words, while CVG is very effective for eliminating duplicates.

TABLE II
FAKE NEWS ASSESSMENT TASK

Method	ROUGE-1 ↑	ROUGE-2 ↑	ROUGE-L ↑
RNN [34], [37]	15.55	2.83	13.08
RNN-context [38]	15.85	3.23	14.48
SummaReranker [39]	23.24	1.32	22.72
EKNet	17.58	3.43	15.34
EKNet+PGN	24.87	3.85	23.42
EKNet+PGN+CVG	**27.45**	**4.33**	**25.65**

The results of this evaluation of the model employing ROUGE-1, ROUGE-2, and ROUGE-L are shown in Table 2 for measuring word overlap, significant segment overlaps, and longest common sequence, respectively. The highest scores are produced by the model proposed by this paper, which stands out as a clear advantage over the competition.

2) Fake News Detection Experiment: As shown in Table 1, the proposed EKNet model is evaluated on three different datasets for fake news detection, namely Baidupedia, Real or Fake, and Fake News Detection Dataset. The model's performance is compared against other models such as FastText[40] , TextRNN[41], TextRCNN[41], Transformer[43] , and Dual[42] .

FastText: FastText is a fast and efficient text classification algorithm that uses word embeddings and n-grams to represent text data. In the context of fake news detection, FastText can be used to classify news articles as either real or fake based on their textual content. The evaluation metrics used were precision (P), recall (R), and F1 score.

TextRNN: TextRNN is a recurrent neural network (RNN) architecture designed for modeling sequential data, such as text. It has been applied to a variety of natural language processing tasks, including language modeling, sentiment analysis, and text classification. In the context of fake news detection, TextRNN can be used to capture the sequential nature of news articles and make predictions based on their content.

TextRCNN: TextRCNN combines recurrent and convolutional neural networks to model text data. In the context of fake news detection, TextRCNN can be used to capture both the sequential and contextual information of news articles and make predictions based on their content.

Transformer: Transformer uses self-attention mechanisms to model sequential data, such as text. Transformer can be used to capture the global information of news articles and make predictions based on their content.

Dual: Dual can be used to improve the performance of existing models by leveraging both labeled and unlabeled data.

As shown in the table 1, EKNet outperforms all other models on all three datasets. Specifically, it achieves the highest precision, recall, and F1 score on all three datasets, demonstrating its effectiveness in addressing the challenges of news credibility assessment. The EKNet achieves a precision of 99.5%, a recall of 99.4%, and an F1 score of 99.5% in dataset Baidupedia, which are significantly higher than the baselines.

From Figure 5 and Figure 6, it can be further observed that the f1 score of the FastText model fluctuates greatly

TABLE III
EXPERIMENTAL RESULTS OF EKNET MODEL WITH DIFFERENT MODALITIES

Features	Confusion Matrix [TP,FP,FN,TN]	Precision ↑	Recall (true positive rate) ↑	Miss rate (false negative rate) ↓	F1 score ↑
Text+Entity-Ontology	[625,3,4,369]	**99.5%**	**99.4%**	**0.6%**	**99.4%**
Text	[608,12,15,366]	98.1%	97.6%	2.4%	97.8 %
Entity-Ontology	[584,28,43,346]	95.4%	93.1%	6.8 %	94.2%

↓ lower values indicate better performance

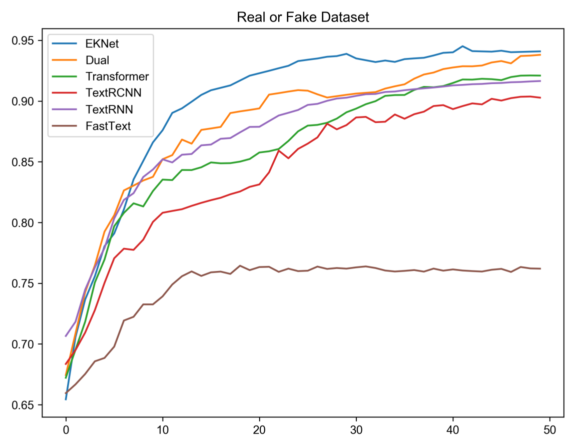

Fig. 5. The performance on each model on Real or Fake Dataset in terms of F1.

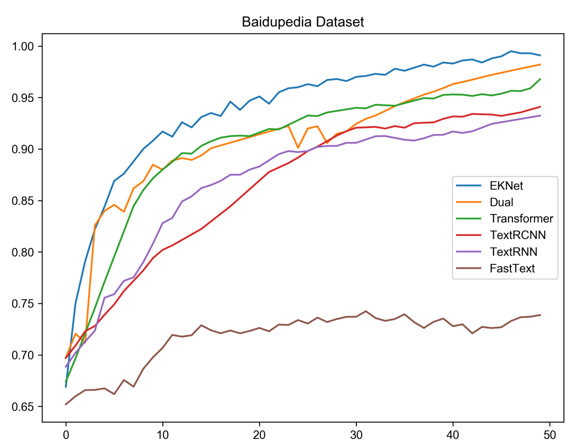

Fig. 6. The performance on each model on Baidupedia Dataset in terms of F1.

during the training process, and the f1 score is relatively low. The performance of other models is better than FastText. On the BaiduPedia dataset, the other models presented by EKnet did not perform well, largely due to the uneven distribution of positive and negative samples in this dataset. The EKNet model effectively utilizes external knowledge bases to enrich news representations, greatly improving the performance of fake news detection.

Overall, the proposed EKNet model shows promising performance in fake news detection compared to other models, especially on the Baidupedia dataset.

E. Ablation Results

To further evaluate the performance of the entity ontology framework and EKNet model, this paper conducted experiments under three different modes: (1) when the input feature was solely the news text, (2) when the input feature was solely the entity ontology, and (3) when both the news text and entity ontology were combined as input features.

The experiments used the following evaluation metrics to measure the performance of the model: Confusion Matrix, Precision, Recall (true positive rate), F1 score and Miss rate (false negative rate). Confusion matrix shows the number of true positives, false positives, true negatives, and false negatives predicted by the model.

The experimental results are shown in table 3. As can be seen from the table, the EKNet model achieved the best performance when both the news text and entity ontology were used as input features. The combination of the two features resulted in a precision of 99.5%, with a recall of 0.994, f1 score of 0.994 and miss rate of 0.006.

In contrast, when the entity ontology was used as the sole input feature, the performance of the model decreased significantly, resulting in a precision of 95.4%, with a recall of 0.931, f1 of 0.978 and miss rate of 0.024.

When the news text was used as the sole input feature, the performance of the model was slightly worse than that of the combined input feature, but still outperformed the entity ontology input feature.

The observed results highlight the critical role of combining both news text and entity ontology features as inputs to the EKNet model for superior performance in news credibility assessment. The entity ontology alone proved insufficient for accurately assessing news credibility, emphasizing the need for complementary information from the news text. While the news text alone provided useful information for assessment, its effectiveness was significantly enhanced when combined with the entity ontology. This underscores the synergistic relationship between textual content and structured ontological knowledge in enhancing the robustness of news credibility assessment models.

VI. CONCLUSION

This paper proposes a novel end-to-end news credibility assessment model, EKNet(Entity Ontology-Based Knowledge Graph Network)2, which is based on knowledge graphs and entity ontology frameworks, aimed at addressing the increasingly serious issue of fake news. The model is experimentally evaluated on multiple datasets, demonstrating significant performance improvements in news credibility assessment. By aligning and fusing entity relation vectors from knowledge graphs with those from news articles, EKNet effectively captures key information in news articles, thus enhancing the accuracy of news credibility assessment. However, the EKNet model still possesses certain potential limitations. For instance, the model's performance might be influenced by the quality and coverage of the knowledge graph. Future research could further explore methods for optimizing knowledge graph construction and incorporating multimodal data into credibility assessment tasks.

REFERENCES

[1] Ma, Jing, Wei Gao, Prasenjit Mitra, Sejeong Kwon, Bernard J. Jansen, Kam-Fai Wong, and Meeyoung Cha."Detecting rumors from microblogs with recurrent neural networks." (2016): 3818.

[2] Conroy, Nadia K., Victoria L. Rubin, and Yimin Chen."Automatic deception detection: Methods for finding fake news." Proceedings of the association for information science and technology 52, no. 1 (2015): 1-4.

[3] Shu, Kai, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu."Fake news detection on social media: A data mining perspective." ACM SIGKDD explorations newsletter 19, no. 1 (2017): 22-36.

[4] Jin, Zhiwei, Juan Cao, Yu-Gang Jiang, and Yongdong Zhang. "News credibility evaluation on microblog with a hierarchical propagation model." In 2014 IEEE international conference on data mining, pp. 230-239. IEEE, 2014.

[5] Ruchansky, Natali, Sungyong Seo, and Yan Liu. "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. 2017.

[6] Ma, Jing, Wei Gao, and Kam-Fai Wong. "Detect rumor and stance jointly by neural multi-task learning." In Companion proceedings of the the web conference 2018, pp. 585-593. 2018.

[7] Shu, Kai, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu."Fake news detection on social media: A data mining perspective." ACM SIGKDD explorations newsletter 19, no. 1 (2017): 22-36.

[8] Jin, Zhiwei, Juan Cao, Yu-Gang Jiang, and Yongdong Zhang. "News credibility evaluation on microblog with a hierarchical propagation model." In 2014 IEEE international conference on data mining, pp. 230-239. IEEE, 2014.

[9] Gupta, Aditi, Ponnurangam Kumaraguru, Carlos Castillo, and Patrick Meier. "Tweetcred: Real-time credibility assessment of content on twitter." In Social Informatics: 6th International Conference, SocInfo 2014, Barcelona, Spain, November 11-13, 2014. Proceedings 6, pp. 228-243. Springer International Publishing, 2014.

[10] Conroy, Nadia K., Victoria L. Rubin, and Yimin Chen. "Automatic deception detection: Methods for finding fake news." Proceedings of the association for information science and technology 52, no. 1 (2015): 1-4.

[11] ElAzab, Ahmed. "Fake accounts detection in twitter based on minimum weighted feature." World (2016).

[12] Yang, Zhiyun, Qi Liu, Hao Wu, Xiaodong Liu, and Yonghong Zhang."CEMA-LSTM: Enhancing contextual feature correlation for radar extrapolation using fine-grained echo datasets." Computer Modeling in Engineering & Sciences 135, no. 1 (2022).

[13] Otter, Daniel W., Julian R. Medina, and Jugal K. Kalita. "A survey of the usages of deep learning for natural language processing." IEEE transactions on neural networks and learning systems 32, no. 2 (2020): 604-624.

[14] Dalpiaz, Fabiano, Alessio Ferrari, Xavier Franch, and Cristina Palomares. "Natural language processing for requirements engineering: The best is yet to come." IEEE software 35, no. 5 (2018): 115-119.

[15] Q. Feng, D. He, Z. Liu, H. Wang, Kim-Kwang Raymond Choo. "SecureNLP: A System for Multi-Party Privacy-Preserving Natural Language Processing", IEEE Transactions on Information Forensics and Security, 15: 3709-3721, 2020, DOI: 10.1109/TIFS.2020.2997134.

[16] Feng, Qi, Debiao He, Zhe Liu, Huaqun Wang, and Kim-Kwang Raymond Choo."SecureNLP: A system for multi-party privacy-preserving natural language processing." IEEE Transactions on Information Forensics and Security 15 (2020): 3709-3721.

[17] Silverman, Greg M., Himanshu S. Sahoo, Nicholas E. Ingraham, Monica Lupei, Michael A. Puskarich, Michael Usher, James Dries et al. "Nlp methods for extraction of symptoms from unstructured data for use in prognostic covid-19 analytic models." Journal of Artificial Intelligence Research 72 (2021): 429-474.

[18] McNamara, Danielle S., Scott A. Crossley, and Rod Roscoe. "Natural language processing in an intelligent writing strategy tutoring system." Behavior research methods 45 (2013): 499-515.

[19] Forestiero, Agostino, and Giuseppe Papuzzo. "Natural language processing approach for distributed health data management." In 2020 28th Euromicro International Conference on Parallel, Distributed and Network-Based Processing (PDP), pp. 360-363. IEEE, 2020.

[20] Potthast, Martin, Johannes Kiesel, Kevin Reinartz, Janek Bevendorff, and Benno Stein. "A stylometric inquiry into hyperpartisan and fake news." arXiv preprint arXiv:1702.05638 (2017).

[21] Popat, Kashyap, Subhabrata Mukherjee, Jannik Strötgen, and Gerhard Weikum. "Credibility assessment of textual claims on the web." In Proceedings of the 25th ACM international on conference on information and knowledge management, pp. 2173-2178. 2016.

[22] Peters, Matthew E., Mark Neumann, Robert L. Logan IV, Roy Schwartz, Vidur Joshi, Sameer Singh, and Noah A. Smith. "Knowledge Enhanced Contextual Word Representations."

[23] Tian, Hao, Can Gao, Xinyan Xiao, Hao Liu, Bolei He, Hua Wu, Haifeng Wang, and Feng Wu. "SKEP: Sentiment Knowledge Enhanced Pre-training for Sentiment Analysis." In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, pp. 4067-4076. 2020.

[24] Bhatt, Shreyansh, Amit Sheth, Valerie Shalin, and Jinjin Zhao. "Knowledge graph semantic enhancement of input data for improving AI." IEEE Internet Computing 24, no. 2 (2020): 66-72.

[25] Zhang, Dehai, Xiaobo Yang, Linan Liu, and Qing Liu. "A knowledge graph-enhanced attention aggregation network for making recommendations." Applied Sciences 11, no. 21 (2021): 10432.

[26] Sang, Lei, Min Xu, Shengsheng Qian, and Xindong Wu. "Knowledge graph enhanced neural collaborative recommendation." Expert systems with applications 164 (2021): 113992.

[27] Hu, Linmei, Mengmei Zhang, Shaohua Li, Jinghan Shi, Chuan Shi, Cheng Yang, and Zhiyuan Liu. "Text-graph enhanced knowledge graph representation learning." Frontiers in Artificial Intelligence 4 (2021): 697856.

[28] Jiang, Zhixue, Chengying Chi, and Yunyun Zhan. "Research on medical question answering system based on knowledge graph." IEEE Access 9 (2021): 21094-21101.

[29] Jiang, Gongyao, Shuang Liu, Yu Zhao, Yueheng Sun, and Meishan Zhang. "Fake news detection via knowledgeable prompt learning." Information Processing & Management 59, no. 5 (2022): 103029.

[30] Liao, Hao, Qixin Liu, Kai Shu, and Xing Xie. "Incorporating user-comment graph for fake news detection." arXiv preprint arXiv:2011.01579 (2020).

[31] Yuan, Xiaoyong, Pan He, Qile Zhu, and Xiaolin Li. "Adversarial examples: Attacks and defenses for deep learning." IEEE transactions on neural networks and learning systems 30, no. 9 (2019): 2805-2824.

[32] Wu, Hao, Qi Liu, and Xiaodong Liu. "A review on deep learning approaches to image classification and object segmentation." Computers, Materials & Continua 60, no. 2 (2019).

[33] Liu, Qi, Jing Zhang, Xiaodong Liu, Yonghong Zhang, Xiaolong Xu, Mohammad Khosravi, and Muhammad Bilal. "Improving wireless indoor non-intrusive load disaggregation using attention-based deep learning networks." Physical Communication 51 (2022): 101584. Harvard

[34] Tai, Yu, Hui He, WeiZhe Zhang, and Yanguo Jia. "Automatic generation of review content in specific domain of social network based on RNN." In 2018 IEEE Third International Conference on Data Science in Cyberspace (DSC), pp. 601-608. IEEE, 2018.

[35] " real_or_fake", 2018, Available: https://www.kaggle.com/rchitic17/real-or-fake.

[36] " Fake News Detection", 2017, Available: https://www.kaggle.com/jruvika/fake-news-detection.

[37] Sinha, Aakash, Abhishek Yadav, and Akshay Gahlot. "Extractive text summarization using neural networks." arXiv preprint arXiv:1802.10137 (2018).

[38] Majid, Rifqi, and Heru Agus Santoso. "Conversations Sentiment and Intent Categorization Using Context RNN for Emotion Recognition." In 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS), vol. 1, pp. 46-50. IEEE, 2021.

[39] Ravaut, Mathieu, Shafiq Joty, and Nancy Chen. "SummaReranker: A Multi-Task Mixture-of-Experts Re-ranking Framework for Abstractive Summarization." In Proceedings of the 60th Annual Meeting of the

Association for Computational Linguistics (Volume 1: Long Papers), pp. 4504-4524. 2022.
[40] Joulin, Armand, Édouard Grave, Piotr Bojanowski, and Tomáš Mikolov. "Bag of Tricks for Efficient Text Classification." In Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers, pp. 427-431. 2017.
[41] Lai, Siwei, Liheng Xu, Kang Liu, and Jun Zhao. "Recurrent convolutional neural networks for text classification." In Proceedings of the AAAI conference on artificial intelligence, vol. 29, no. 1. 2015.
[42] Chen, Qianben, Richong Zhang, Yaowei Zheng, and Yongyi Mao. "Dual contrastive learning: Text classification via label-aware data augmentation." arXiv preprint arXiv:2201.08702 (2022).
[43] Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. "Attention is all you need." Advances in neural information processing systems 30 (2017).

About The Authors

Qi Liu (M'11, SM'18) received his Ph.D. in Data Telecommunications and Networks from the University of Salford, UK in 2006 and 2010. His research interests include Weather Disaster Monitoring and Warning, Edge-Cloud Collaboration, IoT Continuum, and Smart Grid. His recent research esteem covers but is not limited to Weather Radar Extrapolation, NILM Recognition and Prediction, Distributed Resources Scheduling and Efficiency, Short-term Time-series Data Analytics, etc.

Yuan yuan Jin received her B.E. degree in Software Engineering and Qingdao Institute of Technology, China in 2021. And she is currently studying for a M.S. degree in Computer Science and Technology at Nanjing University of Information Science and Technology, China. Her recent research esteem is Natural Language Processing.

Xuefei Cao received her Ph.D. in Information and Telecommunication Engineering from Xidian University, Xi'an, China in 2008. Her research interests include Information Security, Wireless Communication Security and Network Intrusion Detection. Her recent research esteem covers but is not limited to Information Security, AI Security, Vehicular Networks Security, Blockchain, etc.

Xiaodong Liu (M'00, SM'17) received his PhD in Computer Science from De Montfort University and joined Napier in 1999. He is a Reader and is currently leading the Software Systems research group in the IIDI, Edinburgh Napier University. He was the director of Centre for Information & Software Systems. He is an active researcher in software engineering with internationally excellent reputation and leading expertise in context-aware adaptive services, service evolution, mobile clouds, pervasive computing, software reuse, and green software engineering.

Xiaokang Zhou (M'12) is currently an associate professor with the Faculty of Data Science, Shiga University, Japan. He received the Ph.D. degree in human sciences from Waseda University, Japan, in 2014. From 2012 to 2015, he was a research associate with the Faculty of Human Sciences, Waseda University, Japan. He also works as a visiting researcher in the RIKEN Center for Advanced Intelligence Project (AIP), RIKEN, Japan, since 2017. Dr. Zhou has been engaged in interdisciplinary research works in the fields of computer science and engineering, information systems, and social and human informatics. His recent research interests include ubiquitous computing, big data, machine learning, behavior and cognitive informatics, cyber-physical-social systems, cyber intelligence and security. Dr. Zhou is a member of the IEEE CS, and ACM, USA, IPSJ, and JSAI, Japan, and CCF, China.

Yonghong Zhang received his Ph.D. degree from Shanghai Jiao Tong University, China in 2005. His research interests include meteorological disaster monitoring and warning, edge cloud cooperation, IoT continuum, and smart grid. His recent research interests include but are not limited to pattern recognition and intelligent systems, remote sensing big data analysis and deep learning, intelligent equipment and IoT system integration, higher education teaching management and and research work.

Xiaolong Xu received the Ph.D. degree in computer science and technology from Nanjing University, Nanjing, China, in 2016. He was a Research Scholar with Michigan State University, East Lansing, MI, USA, from 2017 to 2018. He is currently a full Professor with the School of Computer and Software, Nanjing University of Information Science and Technology, Nanjing, China. He has authored or coauthored more than 100 peer-review articles in international journals and conferences, including IEEE TITS, IEEE TII, ACM TOIT, ACM TOMM, IEEE IOT, IEEE TCC, IEEE TBD, IEEE TCSS, IEEE TETCI, IEEE ICWS, ICSOC, etc. His research interests include edge computing, the Internet of Things (IoT), cloud computing, and big data. Prof. Xu is a fellow of EAI (European Alliance for Innovation). He was the recipient of the Best Paper Award from the CBD 2016, IEEE CPSCom 2020, and SPDE 2020, the distinguish paper award from EAI Cloudcomp 2019, the best student paper award from EAI Cloudcomp 2019, and the Best Session Paper Award from IEEE DSAA 2020.

Lianyong Qi received the Ph.D. degree from the Department of Computer Science and Technology, Nanjing University, Nanjing, China, in 2011. In 2010, he visited the Department of Information and Communication Technology, Swinburne University of Technology, Melbourne, VIC, Australia. He is currently a Professor with the School of Information Science and Engineering, Qufu Normal University, Rizhao, China. He has chaired two NSFC projects and has published over 40 research papers in international journals, such as the IEEE Transactions on Cloud Computing, the IEEE Transactions on Big Data, the IEEE Transactions on Computational Social Systems, and the IEEE Journal on Selected Areas in Communications and international conferences, such as ICWS, ICSOC, HPCC, and TRUSTCOM. His research interests include big data and services computing.