Analyzing Fake Sports News Detection Methods Using Attention Mechanism and Neural Networks in the Sports Industry

Muqing Zhu[1*]

Abstract

In an era marked by the rapid proliferation of Internet-based social platforms, efficient data dissemination across networks has reached unprecedented levels. However, this widespread accessibility to information has also ushered in a concerning trend—the propagation of false and deceptive narratives. The consequences of fake news, especially in the context of the sports sector, are far-reaching. Beyond eroding trust in media sources, fake sports news can instigate social unrest and disruption with implications extending into politics and the economy.Recognizing the urgent need to address this challenge, this study embarks on the vital task of automatically detecting fake sports news within the complex landscape of online content. Fake sports news typically comprises statements or reports containing elements of falsehood, often diverging significantly from the actual events they purport to describe. Such misinformation is frequently disseminated for political or economic reasons, making its accurate identification formidable.To tackle this pressing issue, we propose a comprehensive fake sports news detection framework termed MFNDF (Multimodal Fake News Detection in Sports). This innovative framework leverages a multifaceted approach, extracting features from three distinct modalities.Firstly, we employ the BERT model for text feature extraction, fine-tuning the extracted text features through a fully connected (FC) layer to enhance the representation of news semantics.Secondly, in image feature extraction, we harness the power of the DenseNet pre-training model to extract convolution features from image content. Additionally, we utilize the Discrete Cosine Transform (DCT) algorithm to extract image frequency domain features, which aid in detecting image tampering and repeated compression—a common tactic employed in creating fake sports news images.Thirdly, we delve into the realm of user context, mining valuable insights from users' behavioral features and news statistical features through advanced feature engineering techniques.Furthermore, we introduce an attention mechanism, a critical component of our approach, to assign weights to word vectors within the text feature space. This attention mechanism takes into account both image features and word vectors, thereby facilitating the fusion of image and text information. The resulting feature vector, enriched with multimodal data, enhances the overall performance of our detection model.Systematic experiments conducted on MFNDF confirm its superior effectiveness in the realm of fake sports news detection. By addressing the critical challenge of identifying deceptive narratives within the sports sector, our research strives to mitigate the adverse impact of fake sports news and safeguard the integrity of information in the dynamic world of sports reporting."

Keywords: Fake sports news detection; Multimodal feature; Attention; Neural network; sports sector

1. Introduction

The digital age has ushered in an unprecedented era of information dissemination, with the internet and social media platforms serving as primary conduits for news and content sharing. However, this proliferation of information has given rise to a significant challenge—fake news. Deceptive narratives, falsified images, and misleading external cues have made identifying fake news a critical concern, particularly within the sports sector (Sahoo & Gupta, 2021; Zhou & Zafarani, 2019). In recent years, fake news has evolved beyond simple textual fabrications, encompassing a broader spectrum of deceptive tactics. While traditional fake news detection models relied predominantly on linguistic characteristics, these approaches often fell short in addressing the complexities of the modern information landscape. This is particularly evident within the realm of sports reporting, where news comprises not only textual content but also visual elements and external information that may subtly hint at deception. Models solely based on unimodal data, such as text or images, have proven inadequate for effectively distinguishing fake sports news. The limitations of such unimodal models include their inability to harness valuable external information and their struggle to capture the nuanced semantics of short text descriptions (Kwon et al., 2013; Zhou & Zafarani, 2020).

Recognizing the multifaceted nature of the challenge, the academic community has increasingly turned its attention to developing comprehensive multimodal fake news detection frameworks. These frameworks draw insights

[1] The University of Sydney, Camperdown NSW 2006, Australia
*Corresponding author:Muqing Zhu, Email:z15162227259@163.com

from various data modalities, including text, images, and external cues, to create a holistic approach to fake news detection.

In response to this pressing need, this study introduces the Multimodal Fake News Detection Framework (MFNDF) tailored specifically for the sports sector. MFNDF leverages the power of advanced natural language processing and computer vision techniques to extract features from multiple data modalities. By integrating information from text, images, and external cues, this framework aims to provide a more robust and accurate means of identifying fake news within the dynamic and fast-paced world of sports reporting. Through a detailed exploration of MFNDF's design and methodology, this research seeks to contribute to the ongoing efforts to combat the proliferation of fake sports news. By developing a multifaceted approach that accounts for the diverse forms of deceptive content within the sports sector, this study endeavors to enhance the integrity of sports news reporting in an era characterized by digital misinformation (Choudhary & Arora, 2021; Kaliyar et al., 2020). This in turn leads to behaviors that infringe on individual rights or endanger society. The dissemination of false information, whether by an individual or the media, can have far-reaching consequences for public peace and stability. When people hear nothing but lies from those in power, they lose faith in those in charge, the truth gets diluted, and social belief systems might even crumble. Therefore, studying how to identify and counteract the spread of sports fake news on online social media has been a topic of intense interest (Faustini & Covoes, 2020; Reis et al., 2019). Fake sports news, according to studies on the dissemination of disinformation, travels far further, more quickly, and to a much wider audience than the truth. Misleading information will negatively affect both individuals and society, and the deliberate propagation of fake sports news may potentially lead to violent situations and harm public safety. The detection of fake news is crucial because it has the potential to mitigate or perhaps eliminate the problem of false information (Cao et al., 2020; Tschiatschek et al., 2018).

Fake sports news prevalent on the Internet and its serious negative effects have become a matter of public concern. At present, there are laws and regulations aimed at infringing false reports to regulate news reports. As vital as it is to hold those responsible for the damage caused by fake sports news to account, it is even more crucial to catch such stories before they can proliferate (Lin et al., 2020; Raja et al., 2020). One step in cleaning up the internet is figuring out how to spot fake sports news as soon as it appears in the media and how to limit the damage that can come from spreading misinformation on social media. Traditional methods of manually detecting fake news are inadequate in light of the modern information boom in online media. Validating the veracity of news articles is a laborious process that requires significant time and effort from relevant professionals. As the need arises, technologies like machine learning, deep learning, and natural language processing can be used to automatically detect fake sports news. Text is the primary format for distributing news in online social network media; images, videos, and other forms of media are sometimes included as well. Most existing techniques for detecting fake sports news go backwards from the news itself, studying the syntax and language style of the news article to pinpoint the source of the fakery. Also, some researchers use the history of news organizations as a secondary identifier. As a result, there has been a surge of interest in the study of fake news identification as a way to lessen the public's exposure to and the news ecosystem's overall vulnerability to this type of content. Detection based on news content alone is becoming increasingly challenging, thus it is essential to fully leverage the external information of news (Song et al., 2021b). This research aims to automatically detect and classify news from beginning to finish using artificial intelligence technology and knowledge of the social network domain to address the aforementioned issues and improve the detection effect of fake news. From a computer science point of view, the major goal of this study is to address the difficulties associated with the fact that conventional methods of detecting false information rely heavily on human intervention and produce slow results. This makes use of deep learning and associated technologies to create diagnostic models for identifying fake sports news. To limit and prevent the spread of false news on social media, it is important to identify such stories as soon as possible after they have been published (Liang et al., 2015; Liu & Xu, 2016).

This work designs a fake sports news detection framework called MFNDF, which extracts features from three modalities. The text feature extraction is based on the BERT model, and the extracted text features are fine-tuned through FC layer to better represent semantics of sports news (Nasir et al., 2021). Image feature extraction uses DenseNet pre-training model to extract image content convolution features and uses DCT algorithm to extract image frequency domain features to represent image tampering and repeated compression information. User context statistical features mining users' behavioral features and sports news statistical features based on feature engineering. This work introduces an attention mechanism to assign weights to word vectors in text through image features and word vectors. After weighted averaging, a feature vector fused with image and text information is obtained, which is added to the multimodal feature set, and the model detection effect is improved. In this work, systematic experiments are conducted on MFNDF, results verify the superiority for method (Alexandre, 2018).

2. Related Work

By combining visual input from the frequency and pixel domains, the novel framework Multi-Domain Visual Neural Network was presented to detect bogus news in reference. Automatic detection of fake sports news was proposed in reference (Guo et al., 2022), which involved traversing and extracting a set of general text properties from sports news in order to build a fake information detector, and then feeding that detector sports news text material. In order to identify unreliable sources, the authors of Reference (Kaliyar et al., 2021) created a logic model that incorporates text sentiment features. The author derives the information's veracity by combining a measuring technique for emotional outliers in text with the text itself, yielding a full ranking of the text's erroneous results. An RNN-based deep attention model was suggested in (Ozbay & Alatas, 2020). The model utilizes soft attention within a recurrent neural network to zero in on critical features for capturing the time-varying context of relevant posts. Soft attention mechanisms have been shown to improve the speed and accuracy with which bogus news may be spotted in experiments. Reference (Ma et al., 2015) studied fake sports news detection using social context, and proposed a three-layer relational embedding framework TriFN. The framework also simulates the relationship between news publishers and news and the interaction between users and sports news and achieves a great improvement in classification performance. Reference (Wang et al., 2018) early proposed a new method based on rumor life cycle time series to capture the temporal features of these features. This also applies time series modeling techniques to integrate various social context information, and achieves good results. Reference (Jin et al., 2017) uses the deep learning framework RNN to analyze the comment text of the sports news, and uses the tf-idf algorithm to extract the representative words of a certain period of time as features. Then input to RNN for detection, which achieved remarkable results at the time. Reference (Qazvinian et al., 2011) proposes a fake sports news detection model for machine learning algorithms based on content-based features and investigates the possible improvements that can be achieved by ensemble learning methods such as AdaBoost and Bagging. Literature (Yang et al., 2012) proposed method to build a Bayesian classifier to detect rumors, which retrieved and classified fake sports news by building different Bayesian classifiers. Each Bayesian classifier corresponds to a feature, and the experimental results have high precision and recall. Reference (Shu et al., 2019) utilizes user reporting information to flag users who are potentially posting fake sports news, and the algorithm performs Bayesian inference to detect fake sports news. The algorithm employs posterior sampling to actively weigh exploitation and exploration to understand user reporting accuracy, demonstrating the power of leveraging user information for fake sports news detection. Literature (Thota et al., 2018) provides a user behavior-based rumor detection approach, and examines the user behavior-based rumor identification scheme, including the features of news followers' comments and forwarding behavior, and the experimental results demonstrate its superior performance. The chronological, structural, and linguistic characteristics of fake sports news are investigated in reference (Cataldi et al., 2013), which provides a rumor detection approach based on these characteristics. Following this, a random forest classifier is built, and a novel model for periodic time series is proposed. According to the concept, rumors may rise and fall at different rates at different times. The experimental results show that this approach successfully categorizes rumors. Reference (Chen et al., 2018) uses multimodal content combined with deep neural networks to tackle the issue of fake sports news detection. For efficiently combining textual, visual, and social context information, they suggest an RNN fused with an attention mechanism. Textual and social context for a tweet are combined with an LSTM to create a single representation. Visual features taken from a pre-trained deep CNN are combined with this shared representation. The output of the LSTM at each time step is then used as a neuron-level attention unit to coordinate visual information during fusion. According to reference (Song et al., 2021a), an end-to-end event adversarial neural network is proposed to detect new false sports news events using multimodal aspects of events. The model incorporates a multimodal feature extractor, a fake sports news detector, and an event discriminator. Articles' textual and visual properties can be extracted separately using a multimodal feature extractor. The event discriminator is responsible for identifying commonalities between events while filtering out information that is unique to individual events. Joint training of a multimodal variation auto encoder and fake sports news detector (Singhal et al., 2020) results in the acquisition of common representations for both textual and visual data. This model has three key components: a variation auto encoder, a probabilistic latent variable model, and an optimizer for the bounds on the marginal likelihood of the observed data. The multimodal representation created by the bimodal variation auto encoder is then used by the fake sports news detector to make a determination as to whether or not the article is fake.

3. Designed Method

3.1. Text Vector Representation

Before deep learning became popular, the extraction of text features mainly relied on some statistical methods. Although this method is simple and fast, it is not scientific and comprehensive enough to only rely on word frequency to measure the importance of words in the text. The text usually has few but important words. At the same time, this calculation method ignores the position information and cannot represent the context information of the text. In order to represent the context information in the word vector layer, word2vec came into being. In this paper, the preprocessed fake sports news text is used to train word2vec. First of all, this paper uses the Chinese word segmentation tool jieba to segment the news text and establish a news dictionary. Then, the word vector is obtained by training the text data after word segmentation. This paper uses the training methods of Negative Sampling and Skip-Gram, which can speed up the model training process while ensuring the accuracy of word vectors. This paper uses the mature gesim toolkit in the sports industry to complete the training of word vectors.

For text data, classification tasks often need to convert natural language text into a vector form that the model can learn. At present, most of them are based on model-based word vector representation. After word2vec was proposed, various word vector pre-training methods have blossomed. BERT, a landmark language expression model, refreshed the task index records of many NLP fields shortly after it was proposed, which is an epoch-making progress in the NLP field. BERT is a deep bidirectional language representation model based on the Encoder structure in Transformer. The basic structure is illustrated in Fig. 1.

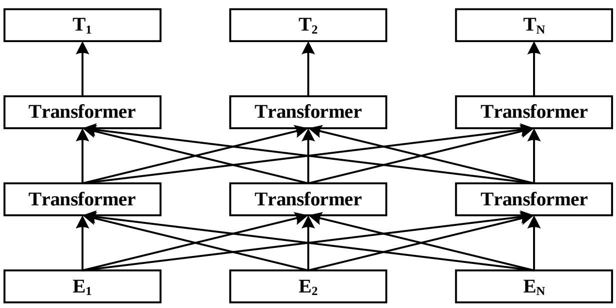

Figure 1. BERT structure.

The essence of the BERT network structure is a stack of Transformers. BERT can accept both single-sentence input and sentence-pair input, depending on the target task. In addition, BERT proposes two different pre-training tasks, MLM and NSP, according to the difference in task objectives. When MLM is pre-trained, some words or characters are randomly masked. MLM cleverly trains the model by predicting masked words, so that each word can well reflect the context structure and represent the context semantics. There are also some problems with this approach, such as the downstream task has no identifier when fine-tuning on the pre-trained model, which leads to inconsistent information in the pre-training and fine-tuning stages. In response to this problem, BERT proposes to randomly replace the identifier with other words. MLM is a trade-off problem. In terms of improved accuracy, the slower convergence speed is acceptable. NSP, as the name suggests, is the next sentence prediction task. Unlike MLM, BERT trains with input sentence pairs. The NSP task expects to discover the relationship between sentences through training, which is suitable for answering systems and natural language reasoning.

In the fake sports news detection framework of this paper, this paper chooses the open source chinese-bert-wwm as the pre-training model of this paper. BERT-wwm is based on the BERT model improvement and improves the model performance by improving the mask strategy. The strategy is to change the word-based mask method in BERT to a phrase-based mask. Compared with the word-based mask method, BERT-wwm uses phrases as tokens. This forces the model to pay attention to more distant context information and has stronger grammar learning ability. When analyzing news data, it is found that in addition to news text content, user descriptions, user locations and news section categories all exist in the form of text. The text content of the news and the user's personal introduction each have a strong influence on the authenticity of the news, and the connection between the two can also be used as an effective feature for discrimination. When the two are correlated, the probability of true news is significantly increased, and when the two are completely uncorrelated, the probability of true news is greatly reduced. Therefore, this article splices the news text with the blogger's self-introduction, and at the same time splices the blogger's geographical information and news section categories into the blogger's introduction and splices the completed text.

$$Text = Test_{news} + Test_{description} + Test_{category} + Test_{location} \quad (1)$$

Where $Test_{news}$ is news feature.

3.2. Image Feature Extraction

With the development of the Internet, the content of images and videos in the news has gradually increased, and the rich pictures often make the news more expressive and make people want to continue reading. Image, an important information source, has important reference significance for fake sports news detection. This section will describe the image feature extraction process of this framework from the perspectives of frequency domain and image content, respectively.

Through data set analysis and research, the pictures in fake sports news are often of lower quality, clarity and resolution than real news pictures. The reason is that fake sports news spreads fast on the Internet. Through direct copying, multiple reposts or tampering and use by communicators, news images are continuously compressed. For the secondary compression and tampering of images, periodic features are displayed in the frequency domain features, which is helpful for the detection and discrimination of fake sports news. Based on the above analysis, the frequency domain feature can represent the compression and tampering information of the image, and it is a good feature with distinguishing degree. In this paper, the DCT algorithm is used to extract the frequency domain features of news images to capture the image structure tampering and compression information. The frequency domain features are extracted from DCT. First, the news image is divided into 64 regions, and DCT transform is performed on each region to extract DCT information. Then, the coefficient information of each frequency is counted by the histogram, and after Fourier transform, a threshold is set to count the number of exceeding the threshold. The last image gets a 64-dimensional statistical feature, which is used as the DCT feature of the image.

Long-term practice shows that the extraction of image features is inseparable from convolutional neural networks. The convolutional neural network greatly reduces the amount of data while retaining the image features to the greatest extent, which is in line with the principles of image processing in this paper. Through the three components of the convolutional neural network, the local feature extraction of images, feature selection and dimensionality reduction, and docking of downstream tasks can be completed. With the emergence of efficient convolutional neural networks, the model has become more efficient and accurate in extracting high-level features from images. This not only relies on massive training data and complex models, but also on powerful computing power. Massive data and high computing power costs have not deterred individual developers from using it. Because most models provide open-source code and provide pre-trained models, individual developers will greatly improve their downstream tasks by virtue of the excellent feature extraction capabilities of pre-trained models. Based on this, this paper chooses to use the pre-trained model as the method for feature extraction of news image content. For the fake sports news dataset, the analysis found that the size of news images is mostly around 512*512. Therefore, first adjust the image size to 512*512 and input it into the pre-trained model Densenet-121. In this paper, the output of the third Transition Layer is selected as the pre-training feature of the image content. The extraction process of the image convolution feature is demonstrated in Figure 2.

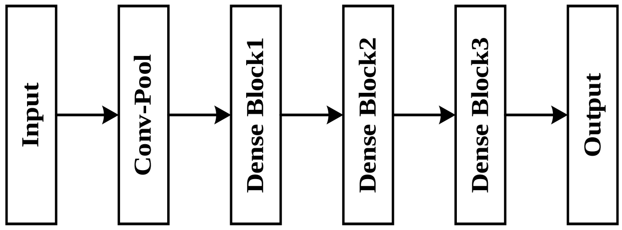

Figure 2. Image convolution feature extraction.

3.3. User Side and Statistical Feature

User state information includes the user's gender, number of fans, number of followers, location, user description, news category and other fields. This paper constructs effective features by mining the hidden information of these fields. Taking the number of users' followers as an example, fake sports news is mainly distributed by users with less than 10,000 followers. At the same time, as the number of fans increases, the proportion of fake sports news in the total news release gradually decreases. Combined with the actual situation, those with a large number of fans are often certified by Weibo or the official platforms of various media, and the news released is usually true. However, those with a small number of fans are often personal bloggers, who publish or retweet news, which is more casual, resulting in the spread of fake news. In addition, the number of fans and followers of users can also be used to mine a large number of hidden features through numerical operations. At the same time, the geographic location information is too detailed, this paper extracts the provincial description of geographic location through regular expressions and uses provincial units as discrete features. Combined with the above analysis, this paper constructs statistical features from three aspects: user information, text statistics and image statistics. These features will be used as user-side contextual features to participate in the construction of the model.

3.4. Multimodal Fake Sports News Detection Framework

This paper proposes a deep learning-based multimodal news detection framework in Figure 3. The whole framework consists of a multimodal feature extraction module, a multimodal feature fusion module and an output module. Among them, the multimodal feature extraction module includes three sub-modules: the text feature extraction module, which uses the BERT model to extract features such as news text semantics and style. The image feature extraction module includes the content information and frequency domain information of the image. The context feature extraction module includes the extraction of user portrait features and statistical features.

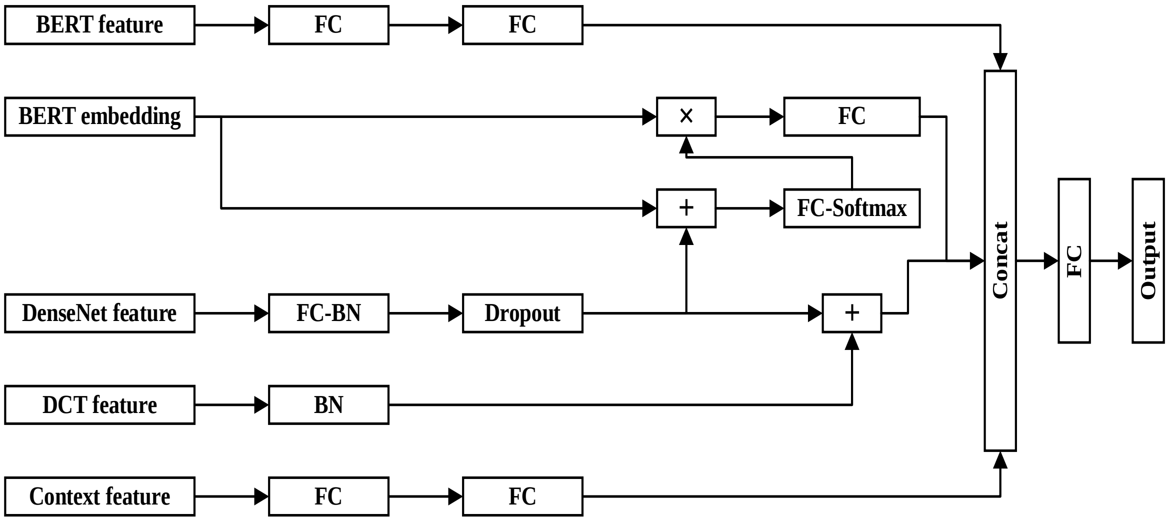

Figure 3. MFNDF architecture.

For text data, it is composed of news text, user geographic location, and user self-description. Generally, BERT can only process text with a length of 512, and text longer than 512 needs to be truncated. This article uses head and tail truncation processing, that is, the first 255 characters and the last 255 characters of the text are preserved as the text input. Then the text segmentation is mapped to the text vector shape through the embedding layer.

$$X_{text}^{i} = [CLS, w_i^{(1)}, w_i^{(2)}, \dots, w_i^{(n)}, SEP] \qquad (2)$$

Where CLS is classification.

In this paper, the text vector of the news is used as the input of the BERT pre-training model, and a vector is obtained by extracting the text features through the BERT pre-training model.

For image data, this paper first adjusts the image to 512*512 pixel size and obtains the input of the image through preprocessing. Select Layer3 in DenseNet121 as the image feature extraction structure, and then connect two fully connected layers. In order to speed up network convergence and prevent overfitting, this paper adds a Batch Normalization layer between the fully connected layer and the activation function. Finally, after Dropout, the obtained vector is the image content feature vector. Frequency domain features are also important as supplementary features of images. For DCT feature extraction, each sample corresponds to a 64-dimensional DCT feature vector. Finally, concat the two features to obtain the image feature vector as the input of the subsequent model.

For user context information, this paper does the following processing. For user age, number of fans, number of followers and other continuous value features with large interval spans, bucketing is used to discretize them. Afterwards, it is converted into a dense vector output through the Dense Embedding layer. The features such as user follower ratio and follower ratio are normalized and input. For the user's region, news types and other information are discrete features, the neural network cannot directly process discrete features, and it needs to be input into the dense embedding layer to convert it into a low-dimensional dense vector. The discrete features input from the user side are spliced with numerical features after passing through Dense Embedding. After two layers of fully connected layers, 256-dimensional user context features are obtained.

$$F_{context} = \tanh\left(w_2 \tanh(w_1 T + b_1) + b_2\right) \quad (3)$$

Where w is weight, b is bias.

News published on the Internet often contains pictures that are closely related to the news content. The way that the simulated person judges the truth of the news is often to first obtain the main content of the news through the keywords of the news, and then observe the key content of the picture. If the picture and text do not match, there is a high probability that it is unreal. Usually, news text and image content are more relevant and more realistic. If the image and text are irrelevant or the difference is small, there is a high probability that it is fake sports news, and you should focus on the words related to the image in the text. This work incorporates an attention mechanism and uses image features and word features to calculate attention weights. Then, the text vector is generated according to the weight, so that the representation of the text feature contains part of the image information, and the image and text features are integrated to a certain extent. The essence of Attention's work can be understood as the weighted average of the vector V according to the vector Q and the vector K to obtain the output vector.

$$ATT(Q,K,V) = \sum Sim(Q,K)V \qquad (4)$$

Where Q is query, K is key, V is value.

The attention mechanism is embodied in the intersection of image features and text features. The basic idea of this improvement is to use the intersection of image and text features to enhance the expression of text features.

Finally, the multimodal feature vector is obtained by concat splicing. The feature vector is input into the feedforward network for classification, and the prediction result is obtained.

$$y_{pred} = Sigmoid(\tanh(w_2 \tanh(w_1 F + b_1) + b_2)) \quad (5)$$

Where w is weight, b is bias.

The goal of this paper is to detect fake sports news through multimodal information, which is essentially a binary classification task. In this paper, the cross entropy loss is selected as the loss function of the model

4. Experiment

To effectively evaluate the effect of the model, the dataset is usually divided into a training set and validation set. But only one division is not enough to learn the distribution of the entire dataset, and the above operations need to be repeated, which is cross-validation. Cross-validation can better evaluate the predictive ability of the model and reduce over-fitting, so this paper uses five-fold cross-validation for the division of the dataset. Each time the sample set is divided into 5 parts, one sub-sample data is used for model validation, and the other 4 sub-samples are used for training, which is repeated 5 times. The software and hardware experimental environment of this paper is illustrated in Table 1. Accuracy and F1 score are evaluation metrics for this work.

$$Accuracy = \frac{TP+TN}{N+P} \qquad (6)$$

$$F1\ score = \frac{2*Pre*Rec}{Pre+Rec} \qquad (7)$$

Table 1

Relevant experimental environment.

Property	Parameter
Processor	AMD RYZEN
System memory	16.0 GB
Operating system	Ubuntu18 04
Gpu	Nvidia RTX 2080 super*4
Video memory	32.0 GB
Experimental environment	Python3.7

This paper first analyzes the training performance of MFNDF. The main analysis objects are the accuracy rate and F1 score in the training process, and the experimental data is illustrated in Fiure. 4.

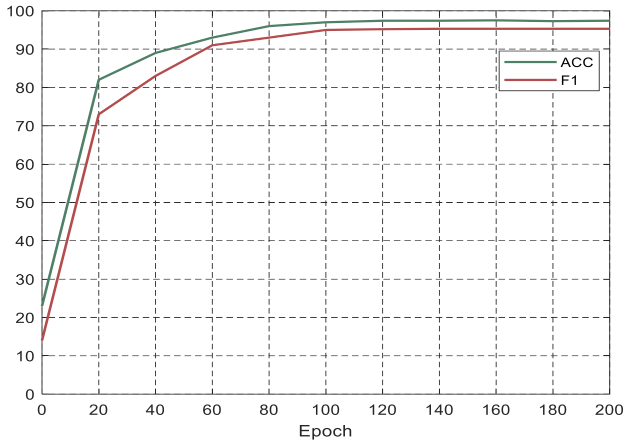

Figure 4. MFNDF training analysis.

As the training progresses, the training accuracy and F1 score of MFNDF gradually increase, and finally the two stabilize and converge.

This work compares MFNDF with other similar methods to test the effectiveness of this method. The comparative experimental results are illustrated in Table 2.

Table 2

Method comparison.

Method	Acc	F1
BERT-Text	93.2	90.2
DenseNet-Image	91.5	88.7
LR-Context	90.5	87.6
MFNDF	95.3	92.2

Compared with other methods, MFNDF can achieve the best performance. This verifies the correctness of the MFNDF design in this paper.

To ensure the balance between the various modal features when designing the network, this paper uniformly sets the features to the same dimension and then splices them into the classifier. To further study the influence of the feature vector dimension on the experimental results, the feature dimension is set to different dimensions for training, and the classification accuracy of the test set is observed in Fig. 5.

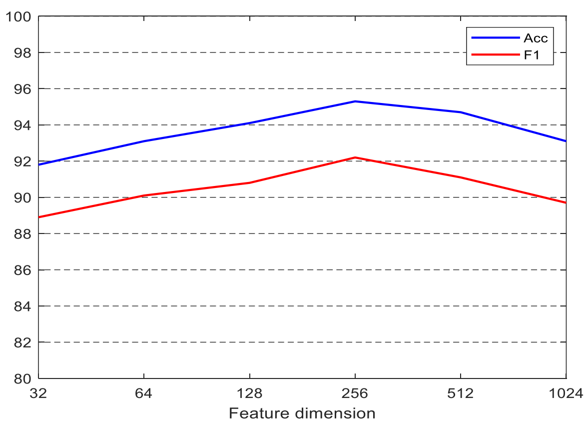

Figure 5. Feature dimension analysis

It is observed that the feature dimension has a certain influence on the experimental effect, and the feature dimension is optimal at 256. According to the analysis of the experimental results, it shows that if the feature dimension is too small, the feature expression ability is not strong, and it cannot carry enough information, resulting in poor model effect. When the feature dimension is too large, the feature will be too sparse, and the feature redundancy and too many parameters will affect the effect and performance of the model.

To further explore the influence of similarity calculation methods in different attention mechanisms on the experimental results. This paper compares existing attention calculations. The existing similarity calculation methods mainly include dot product and scaled dot product. The experimental data is illustrated in Figure 6.

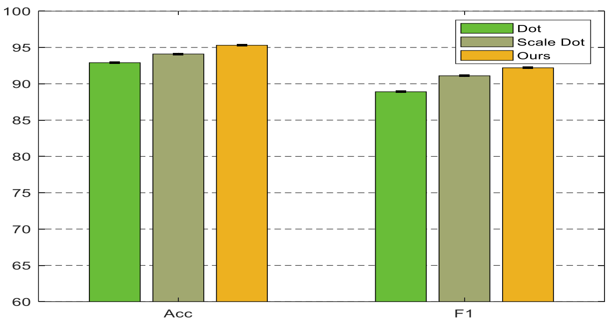

Figure. 6. Attention calculation analysis.

The experimental results show that the calculation effect of using dot product and scaled dot product is slightly worse than the similarity calculation method used in this paper. The reason for the analysis is that the similarity calculation method based on dot product and scaled dot product is too simple and fixed, which is suitable for use in scenarios that require high interpretability. The similarity calculation method used in this paper, because the fully connected layer and softmax are used, there are many learnable parameters in the network. In this way, during model training, the parameter weights will be updated according to the feedback of the learning task, so as to dynamically correct the attention weights.

In MFNDF, BN is used for feature normalization. In order to verify the feasibility of this measure, this work analyzes it, and the experimental data is illustrated in Figure 7.

After processing the features with BN, both the accuracy and F1 score of MFNDF can be improved. This proves the correctness of this strategy.

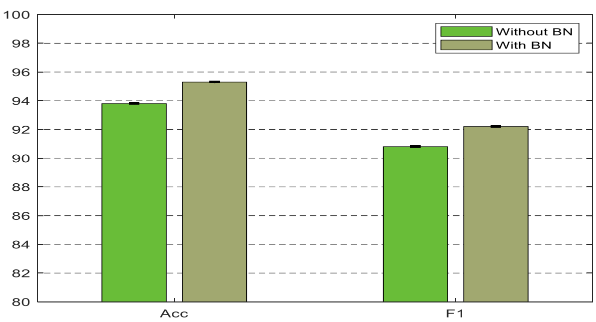

Figure 7. *BN layer analysis.*

5. Conclusion

In an era characterized by the pervasive dissemination of fake news online, the reliance solely on human discernment to identify deceptive content has become insufficient. It is imperative to develop and deploy robust algorithms capable of effectively detecting fake news, especially within the context of the sports sector. The evolution of fake news has seen it take on increasingly diverse forms, encompassing not only false textual and visual content but also external information that subtly implies its deceptive nature. Consequently, conventional fake news detection models, primarily reliant on linguistic cues, face numerous limitations. Firstly, models that exclusively employ unimodal data, such as text or images, tend to exhibit suboptimal performance. They fail to harness the valuable insights embedded in external information that often carries pivotal discriminative features. Secondly, these models struggle to effectively capture the semantic nuances of short text information, which typically possesses sparse semantics.

Lastly, a single modality approach overlooks the intricate feature connections between news images, both at the physical and semantic levels. As a result, a holistic model that comprehensively integrates multimodal features from text, images, and external information has garnered significant attention within the academic community. This research introduces the Multimodal Fake News Detection Framework (MFNDF) tailored specifically for the sports sector. MFNDF adopts a multifaceted approach, extracting features from three essential modalities. For text feature extraction, the model harnesses the power of BERT, fine-tuning the extracted text features through a fully connected (FC) layer, enhancing the representation of news semantics.

In the domain of image feature extraction, DenseNet pre-training models are employed to extract convolution features from image content. Additionally, the Discrete Cosine Transform (DCT) algorithm is utilized to extract image frequency domain features, effectively capturing image tampering and repeated compression—common tactics employed in the creation of deceptive sports news images. Furthermore, the research delves into the realm of user context, leveraging behavioral features from users and statistical features from news through advanced feature engineering techniques. To bolster the model's performance, an attention mechanism is introduced, assigning weights to word vectors within the text feature space based on image features and word vectors. This weighted averaging process yields a feature vector infused with both image and text information, enhancing the overall effectiveness of the detection model. In conclusion, this research addresses the critical challenge of identifying fake news within the sports sector by introducing a comprehensive, multimodal approach. By doing so, it strives to safeguard the integrity of sports news reporting in an age where misinformation abounds.

Data availability statement

The datasets used during the current study are available from the corresponding author on reasonable request.

Conflict of interest

Declares that he has no conflict of interest.

Reference

Alexandre, C. M.-A., Ricardo, V. d.-S., Daniel, A.-B., Dumitriu, Z.-S., & Salvador, J. d.-T. A. A. (2018). Lessons Learned After 366 Thermoablated Veins. *Vascular & Endovascular Review, 1*(1), 22-26. https://doi.org/10.15420/ver.2018.13.1

Cao, J., Qi, P., Sheng, Q., Yang, T., Guo, J., & Li, J. (2020). Exploring the Role of Visual Content in Fake News Detection. In *Disinformation, Misinformation, and Fake News in Social Media: Emerging Research Challenges and Opportunities* (pp. 141-161). Springer. https://doi.org/10.1007/978-3-030-42699-6_8

Cataldi, M., Ballatore, A., Tiddi, I., & Aufaure, M.-A. (2013). Good Location, Terrible Food: Detecting Feature Sentiment in User-generated Reviews. *Social Network Analysis and Mining, 3*(4), 1149-1163. https://doi.org/10.1007/s13278-013-0119-7

Chen, T., Li, X., Yin, H., & Zhang, J. (2018). Call Attention to Rumors: Deep Attention Based Recurrent Neural Networks for Early Rumor Detection. In *Trends and Applications in Knowledge Discovery and Data Mining* (pp. 40-52). Springer. https://doi.org/10.1007/978-3-030-04503-6_4

Choudhary, A., & Arora, A. (2021). Linguistic feature based learning model for fake news detection and classification. *Expert Systems with Applications, 169*, 114171. https://doi.org/10.1016/j.eswa.2020.114171

Faustini, P. H. A., & Covoes, T. F. (2020). Fake news detection in multiple platforms and languages. *Expert Systems with Applications, 158*, 113503. https://doi.org/10.1016/j.eswa.2020.113503

Guo, Z., Shen, Y., Bashir, A. K., Yu, K., & Lin, J. C. w. (2022). Graph embedding-based intelligent industrial decision for complex sewage treatment processes. *International Journal of Intelligent Systems, 37*(12), 10423-10441. https://doi.org/10.1002/int.22540

Jin, Z., Cao, J., Guo, H., Zhang, Y., & Luo, J. (2017). Multimodal Fusion With Recurrent Neural Networks for Rumor Detection on Microblogs. In *Proceedings of the 25th ACM International Conference on Multimedia* (pp. 795-816). ACM. https://doi.org/10.1145/3123266.3123454

Kaliyar, R. K., Goswami, A., & Narang, P. (2021). Fakebert: Fake news detection in social media with a BERT-based deep learning approach. *Multimedia Tools and Applications, 80*(8), 11765-11788. https://doi.org/10.1007/s11042-020-10183-2

Kaliyar, R. K., Goswami, A., Narang, P., & Sinha, S. (2020). Fndnet–a Deep Convolutional Neural Network for Fake News Detection. *Cognitive Systems Research, 61*, 32-44. https://doi.org/10.1016/j.cogsys.2019.12.005

Kwon, S., Cha, M., Jung, K., Chen, W., & Wang, Y. (2013). Prominent Features of Rumor Propagation in Online Social Media. In *2013 IEEE 13th International Conference on Data Mining* (pp. 1103-1108). IEEE. https://doi.org/10.1109/ICDM.2013.61

Liang, G., He, W., Xu, C., Chen, L., & Zeng, J. (2015). Rumor Identification in Microblogging Systems Based on Users' Behavior. *IEEE Transactions on Computational Social Systems, 2*(3), 99-108. https://doi.org/10.1109/TCSS.2016.2517458

Lin, X., Wu, J., Bashir, A. K., Li, J., Yang, W., & Piran, M. J. (2020). Blockchain-based incentive energy-knowledge trading in IoT: Joint power transfer and AI design. *IEEE Internet of Things Journal, 9*(16), 14685-14698. https://doi.org/10.1109/JIOT.2020.3024246

Liu, Y., & Xu, S. (2016). Detecting rumors through modeling information propagation networks in a social media environment. *IEEE Transactions on Computational Social Systems, 3*(2), 46-62. https://doi.org/10.1109/TCSS.2016.2612980

Ma, J., Gao, W., Wei, Z., Lu, Y., & Wong, K.-F. (2015). Detect Rumors Using Time Series of Social Context Information on Microblogging Websites. In *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management* (pp. 1751-1754). ACM. https://doi.org/10.1145/2806416.2806607

Nasir, J. A., Khan, O. S., & Varlamis, I. (2021). Fake News Detection: a Hybrid Cnn-rnn Based Deep Learning Approach. *International Journal of Information Management Data Insights, 1*(1), 100007. https://doi.org/10.1016/j.jjimei.2020.100007

Ozbay, F. A., & Alatas, B. (2020). Fake News Detection Within Online Social Media Using Supervised Artificial Intelligence Algorithms. *Physica a: Statistical Mechanics and Its Applications, 540*, 123174. https://doi.org/10.1016/j.physa.2019.123174

Qazvinian, V., Rosengren, E., Radev, D., & Mei, Q. (2011). Rumor Has It: Identifying Misinformation in Microblogs. In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing* (pp. 1589-1599). Association for Computational Linguistics. https://aclanthology.org/D11-1147

Raja, G., Ganapathisubramaniyan, A., Anbalagan, S., Baskaran, S. B. M., Raja, K., & Bashir, A. K. (2020). Intelligent Reward-based Data Offloading in Next-generation Vehicular Networks. *IEEE Internet of Things Journal, 7*(5), 3747-3758. https://doi.org/10.1109/JIOT.2020.2974631

Reis, J. C., Correia, A., Murai, F., Veloso, A., & Benevenuto, F. (2019). Supervised Learning for Fake News Detection. *IEEE Intelligent Systems, 34*(2), 76-81. https://doi.org/10.1109/MIS.2019.2899143

Sahoo, S. R., & Gupta, B. B. (2021). Multiple Features Based Approach for Automatic Fake News Detection on Social Networks Using Deep Learning. *Applied Soft Computing, 100*, 106983. https://doi.org/10.1016/j.asoc.2020.106983

Shu, K., Wang, S., & Liu, H. (2019). Beyond News Contents: the Role of Social Context for Fake News Detection. In *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining* (pp. 312-320). ACM. https://doi.org/10.1145/3289600.3290994

Singhal, S., Kabra, A., Sharma, M., Shah, R. R., Chakraborty, T., & Kumaraguru, P. (2020). Spotfake+: a Multimodal Framework for Fake News Detection via Transfer Learning (Student Abstract). *Proceedings of the AAAI Conference on Artificial Intelligence, 34*(10), 13915-13916. https://doi.org/10.1609/aaai.v34i10.7230

Song, C., Ning, N., Zhang, Y., & Wu, B. (2021a). A multimodal fake news detection model based on crossmodal attention residual and multichannel convolutional neural networks. *Information Processing & Management, 58*(1), 102437. https://doi.org/10.1016/j.ipm.2020.102437

Song, C., Shu, K., & Wu, B. (2021b). Temporally Evolving Graph Neural Network for Fake News Detection. *Information Processing & Management, 58*(6), 102712. https://doi.org/10.1016/j.ipm.2021.102712

Thota, A., Tilak, P., Ahluwalia, S., & Lohia, N. (2018). Fake News Detection: a Deep Learning Approach. *SMU Data Science Review, 1*(3), 10. https://scholar.smu.edu/datasciencereview/vol1/iss3/10

Tschiatschek, S., Singla, A., Gomez Rodriguez, M., Merchant, A., & Krause, A. (2018). Fake News Detection in Social Networks via Crowd Signals. In *Companion Proceedings of the the Web Conference 2018* (pp. 517-524). ACM. https://doi.org/10.1145/3184558.3188722

Wang, Y., Ma, F., Jin, Z., Yuan, Y., Xun, G., Jha, K., Su, L., & Gao, J. (2018). EANN: Event Adversarial Neural Networks for Multi-modal Fake News Detection. In *Proceedings of the 24th ACM Sigkdd International Conference on Knowledge Discovery & Data Mining* (pp. 849-857). ACM. https://doi.org/10.1145/3219819.3219903

Yang, F., Liu, Y., Yu, X., & Yang, M. (2012). Automatic Detection of Rumor on Sina Weibo. In *Proceedings of the ACM Sigkdd Workshop on Mining Data Semantics* (pp. 1-7). ACM. https://doi.org/10.1145/2350190.2350203

Zhou, X., & Zafarani, R. (2019). Network-based Fake News Detection: a Pattern-driven Approach. *ACM Sigkdd Explorations Newsletter, 21*(2), 48-60. https://doi.org/10.1145/3373464.3373473

Zhou, X., & Zafarani, R. (2020). A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities. *ACM Computing Surveys (CSUR), 53*(5), 1-40. https://doi.org/10.1145/3395046