



DOI: 10.32604/iasc.2023.037599

Article



SNG-TE: Sports News Generation with Text-Editing Model

Qiang Xu*, Wei Zhang, Hui Ding and Shengwei Ji

HeFei University, Hefei, 230031, China

*Corresponding Author: Qiang Xu. Email: xuqiang@hfuu.edu.cn

Received: 10 November 2022; Accepted: 28 February 2023

Abstract: Currently, the amount of sports news is increasing, given the number of sports available. As a result, manually writing sports news requires high labor costs to achieve the intended efficiency. Therefore, it is necessary to develop the automatic generation of sports news. Most available news generation methods mainly rely on real-time commentary sentences, which have the following limitations: (1) unable to select suitable commentary sentences for news generation, and (2) the generated sports news could not accurately describe game events. Therefore, this study proposes a sports news generation with text-editing model (SNG-TE) is proposed to generate sports news, which includes selector and rewriter modules. Within the study context, a weight adjustment mechanism in the selector module is designed to improve the hit rate of important sentences. Furthermore, the text-editing model is introduced in the rewriter module to ensure that the generated news sentences can correctly describe the game events. The annotation and generation experiments are designed to evaluate the developed model. The study results have shown that in the annotation experiment, the accuracy of the sentence annotated by the selector increased by about 8% compared with other methods. Moreover, in the generation experiment, the sports news generated by the rewriter achieved a 49.66 ROUGE-1 score and 21.47 ROUGE-2, both of which are better than the available models. Additionally, the proposed model saved about 15 times the consumption of time. Hence, the proposed model provides better performance in both accuracy and efficiency, which is very suitable for the automatic generation of sports news.

Keywords: Text editing; news generation; automatic generation

1 Introduction

Automatic news generation is an important artificial intelligence direction in journalism, which improves work efficiency and substantially reduces costs. Currently, there are several robots available



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

in the industry, such as Wordsmith¹, Heliograf², and DT draft³. These robots significantly cut labor costs down and improve news generation efficiency.

Many sports games are held on daily bases worldwide. Besides live broadcasts, users can keep abreast of the process through sports news. Large texts are generated by live broadcast that objectively represents the sports games. Therefore, automatic generation is suitable for the creation of sports news.

Previously, multiple studies have focused on the automatic generation of sports news. The rival generation models can mainly be divided into extractive and generative. Extractive models select important sentences from live text to compose sports news [1–5]. However, suitable sentences are not selected because of the selection mechanism. In addition, a human writing style cannot be formed because commentary sentences are used directly. Although the use of artificial templates [6,7] can alleviate these problems, writing templates costs a lot. Generative models mainly use the Seq2Seq model to generate sports news [8,9]. However, the Seq2Seq model may cause hallucinations which the output text contains information that was not presented in the input text [10]. Furthermore, it requires a long duration to generate the texts. As demonstrated in Fig. 1, there are two limitations in sports news generation at present: unable to select the suitable commentary sentence and the generated news sentences with poor quality. Accordingly, this study proposes an approach to sports news generation to solve the above limitations. The main contributions are detailed as follows:

- (1) Designing the selector using an automatic annotation algorithm to improve the hit rate of important sentences.
- (2) Introducing a text-editing model in rewriter, which ameliorates the problem of hallucinations and the efficiency of sports news generation.
- (3) Demonstrating experimental results that indicate the improved accuracy and efficiency of the model in sports news generation.

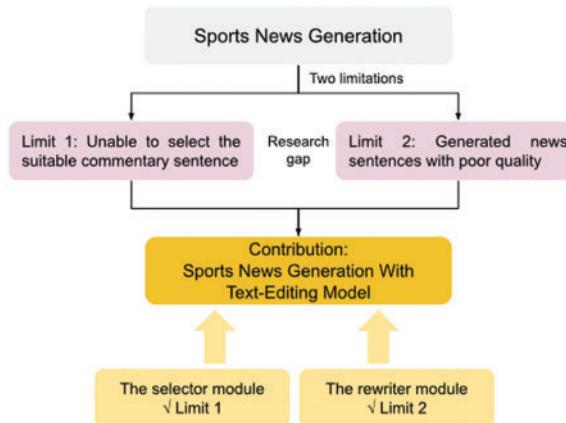


Figure 1: Motivation of sports news generation

2 Related Work

Many methods have been proposed and applied to the field of journalism. For instance, in detecting fake news, Kiran et al. [11] introduced a novel multi-stage transfer learning approach

¹<https://automatedinsights.com/wordsmith/>

²https://www.washingtonpost.com/allmetsports/2017-fall/games/football/87055/?utm_term=.1532c57a69dd&itid=lk_inline_manual_4

³<https://www.yicai.com/author/794.html>

for effective fake news detection, where it utilizes users' comments as auxiliary information. Zakharchenko et al. [12] studied whether newsmakers could manipulate the media with fake news. Alzhrani [13] used various neural network structures, including text convolutional neural network (TextCNN) and hierarchical attention networks (HAN), to detect political ideology in the news. Hnaif et al. [14] utilized various classifiers such as support vector machine and naive Bayes to analyze sentiment tendencies in the news. Oh et al. [15] proposed an automated news generation system, namely News-robot, which can automatically generate news in real-time and enrich user experience by providing users with individualized news according to their choices while expressing various multimedia modalities.

In sports journalism, the Chinese dataset of sports news generation was first released by Zhang et al. [3], which contains 150 sports game data. Thereafter, Huang et al. [8] proposed SportsSum, which collected data on 5428 games in 7 famous leagues and removed the hypertext markup language (HTML) tags and background introductions at the beginning of the news. On this basis, Wang et al. [16] cleaned up the noise data and released SportsSum2.0. Later, Wang et al. [9] expanded the number of samples and released the K-SportsSum database, which contains 7854 game data. Additionally, the latest work by Wang et al. [17] released the first English competition data set goal, which contains 103 samples.

Zhang et al. [3] proposed using the learning ranking model to select important commentary sentences. Zhu et al. [5] selected sentences through conditional random fields. Yao et al. [2] considered the sports news' writing style and the importance of the described events when selecting important sentences. Tang et al. [1,4] identified important sentences to form sports news by building a keywords dictionary. Liu et al. [6,7] used artificial templates to assist news generation. Huang et al. [8] proposed a two-step generation framework by extracting important sentences and then using the Seq2Seq model to rewrite the sports news. On this basis, Wang et al. [9,16] selected the generated sentences through an improved maximal marginal relevance (MMR) algorithm to improve the quality of the sports news.

3 Automatic Generation of Sports News

The automatic generation task and the framework of SNG-TE are introduced in Sections 3.1 and 3.2. The sentence mapping process is described in Section 3.3. Finally, the selector and rewriter modules are introduced in Sections 3.4 and 3.5, respectively.

3.1 Task Formulation

As shown in Fig. 2, the task of automatic generation aims to extract important sentences from a live broadcast commentary set $C = \{(t_1, c_1, s_1), \dots, (t_n, c_n, s_n)\}$ to generate a piece of sports news $R = \{r_1, r_2, \dots, r_m\}$. (t_n, c_n, s_n) represents any n -th commentary, where t is the timeline feature, c is the commentary sentence, s is the current score, and r_m is m -th news sentence. Generated sports news R should cover as many important events in the game as possible.

3.2 Framework

As illustrated in Fig. 3, our work is divided into the following three steps. First, each commentary sentence is mapped to a news sentence. After that, the mapped sentences are automatically annotated through the proposed algorithm used to train the selector to extract important commentary. Finally, the mapped and annotated sentences are combined to form sentence pairs <commentary, news sentence> according to timeline information. These sentence pairs are utilized for training text-editing model to convert commentary into sports news.

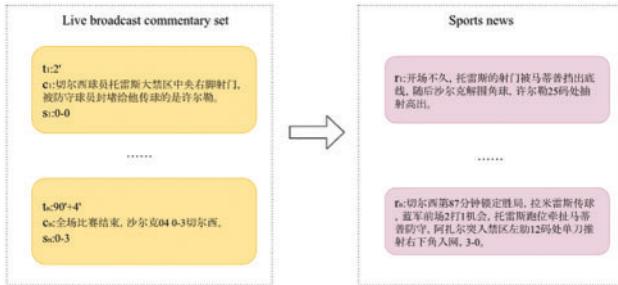


Figure 2: The task of automatic generation is to extract important sentences from a live broadcast commentary set and generate sports news

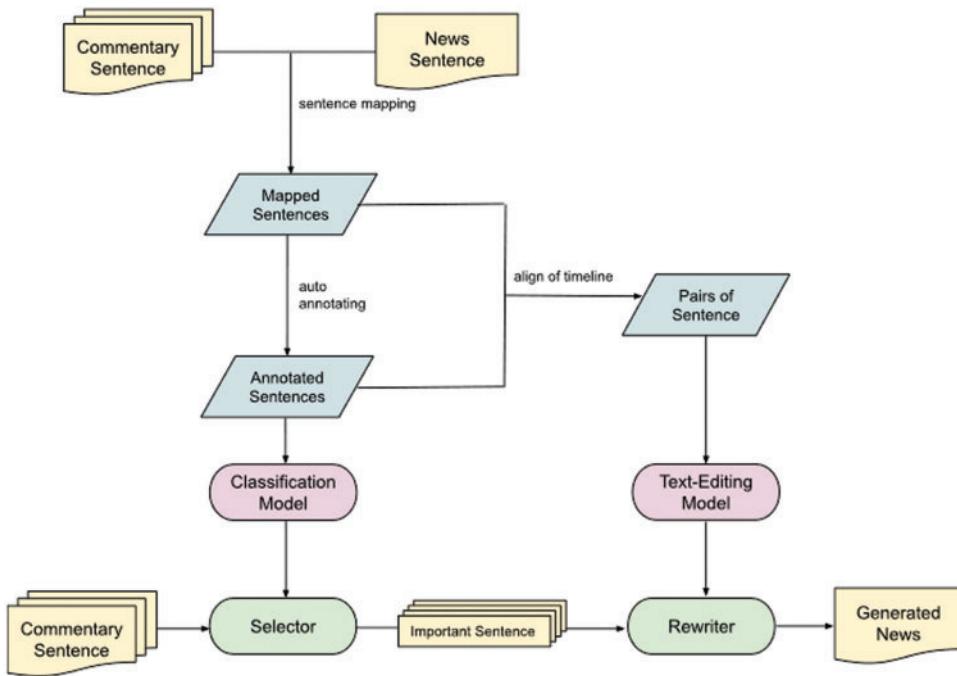


Figure 3: SNG-TE framework process. The sentences should be preprocessed, then used to train classification and text-editing models. The classification model is a selector to extract important sentences, and the text-editing model is a rewriter to convert important sentences into sports news

3.3 Sentence Mapping

The sentences should be annotated to train the selector. It can be noticed that most news sentences start with “in the i -th minutes.”, these timeline information was extracted from each news sentence. Thereafter, the commentary sentences from the time period $i - 2$ to $i + 2$ are extracted.

This work was divided into the following two steps: (1) extracting the timeline feature of each news sentence and obtaining a news set $\text{News} = \{(h_1, n_1), \dots, (h_i, n_i)\}$, where (h_i, n_i) is the news sentence n_i in the h_i minutes; (2) aggregating the commentary c_m with t_i , which from h_{i-2} to h_{i+2} , composed the commentary set $C_{\text{candidate}} = \{(t_1, (c_1, \dots, c_m)), \dots, (t_i, (c_1, \dots, c_m))\}$.

The mapped sentences were automatically annotated to indicate whether they are important (Section 3.4.2). After that, these annotated sentences were used to train the selector. In addition, the sentence annotated as labels ‘1’ were combined with the corresponding news sentence into sentence pairs for training rewriter, as shown in Fig. 4.

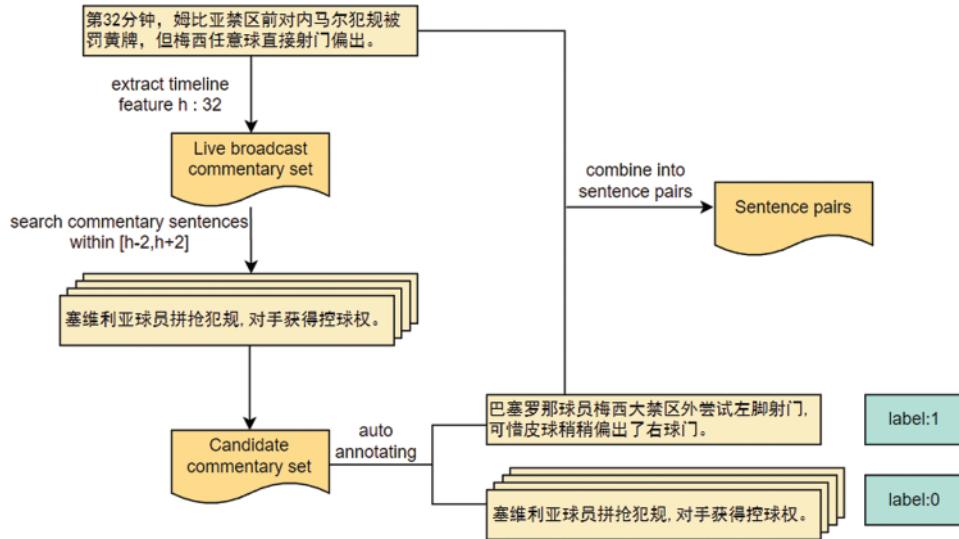


Figure 4: A case of sentence mapping. The timeline feature of each news sentence is extracted, then the commentary sentences are aggregated from the live broadcast commentary set to compose the candidate commentary set. The sentences in the candidate commentary set are annotated for the selector. The sentences annotated label ‘1’ are combined with the corresponding news sentence into pairs for the rewriter

3.4 Selector

The selector is a classification model used to extract important sentences. Particularly, an automatic annotation algorithm is proposed to obtain high-quality training data and save the cost of annotation.

3.4.1 Task Formulation

Given the mapped sentence $C_{candidate} = \{(t_i, (c_1, \dots, c_m)), \dots, (t_i, (c_1, \dots, c_m))\}$, all sentences c need to be annotated to train the selector to extract important commentary.

3.4.2 Automatic Annotation Algorithm

A live broadcast can have thousands of commentary, and manual annotation costs a lot. Therefore, an automatic annotation algorithm is designed to replace manual annotation.

The proposed algorithm is based on Sentence-Bert (SBERT) [18] and KEYWORDS to score and annotate sentences with different labels. The process of automatic annotation is demonstrated in Fig. 5: (1) Calculating similarity score and keywords score and then adding the two scores according to the weight adjustment mechanism. (2) Giving the sentence with the highest score in each group labeled as “1” and the others labeled “0” (3) During the annotating process, a sentence may be annotated more than once. If there are different results, the default is to annotate the sentence with the label “1”.

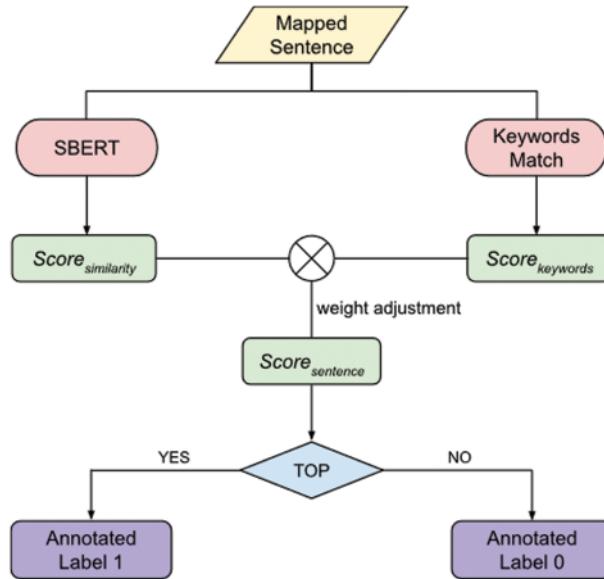


Figure 5: Automatic annotation process. Each mapped sentence's similarity and keyword match scores are calculated and added through the weight adjustment mechanism. Finally, the sentences with the highest scores in each group are labeled as selected and the others as unselected

For $C_{candidate} = \{(t_i, (c_1, \dots, c_m)), \dots, (t_i, (c_1, \dots, c_m))\}$ and $News = \{(h_i, n_i), \dots, (h_i, n_i)\}$, the similarity between commentary c_m and news sentence n_i are calculated through SBERT. The sentences (c_1, \dots, c_m) and n_i are input into SBERT. The sentences are mapped to a 512-dimensional dense vector space and $Score_{similarity}$ is obtained by calculating the cosine value.

Based on the sports news characteristics, keywords such as shooting, foul, and goal are often included in the description of some important sports events. However, they are not considered in the similarity calculation process. A keyword matching is added to solve this problem to improve annotation accuracy.

The top 150 verbs were extracted from the sports news, and the words that were not important were manually removed. Thereafter, these verbs were classified into the following categories: shooting, missing, passing, blocking, defense, foul, and goal. Moreover, different scores were assigned based on the importance of various keywords to the sports games, as shown in Table 1. Each sentence was matched with a regular, and $Score_{keywords}$ is obtained based on the category of keywords contained in the sentence.

Table 1: List of keywords. There are seven verb categories, and some examples are listed in the table. The corresponding category score will be obtained if these keywords are used in the commentary

Categories of keywords	Keywords	Score
Shooting	射门、打门、吊射、推射、头球.....	0.2
Missed shot	偏出、打偏、打飞、顶高、蹭出.....	0.1
Passing	传中、传球、直塞、挑传、横敲.....	0.1

(Continued)

Table 1: Continued

Categories of keywords	Keywords	Score
Blocking	扑出、没收、托出、救下、抱住.....	0.1
Defense	解围、破坏、铲除、化解.....	0.1
Foul	犯规、警告、判罚、领到、出示.....	0.2
Goal	进了、进球	0.5

Finally, the $Score_{sentence}$ not only includes $Score_{similarity}$, but also adds $Score_{keywords}$ as rewards. This Eq. (1) adjusts the weight between them via a set of parameters. After many iterations, the optimized parameters were found as $\alpha = 0.7$ and $\beta = 0.35$.

$$Score_{sentence} = \alpha \times Score_{similarity} + \beta \times Score_{keywords} \quad (1)$$

3.4.3 Classification Model

In this section, a simple classification model was built. The bidirectional encoder representations from transformer (BERT) were used as a pre-trained language model [19], and a linear layer was added at its end to perform the classification task.

BERT is a pre-trained language model built on the bidirectional transformer encoder. The transformer encoder is based on a self-attention mechanism. The BERT is divided into two steps: pre-training and fine-tuning. In the pre-training stage, the BERT adopted two self-supervised techniques: masked language modeling (MLM) and next sentence prediction (NSP) [20]. In the MLM, the BERT randomly predicted the masked input tokens, whereas, in the NSP, the BERT predicted whether two input sentences were sequential. For fine-tuning, the BERT was initialized with the pre-trained parameter weights, and all parameters were fine-tuned using labeled data from downstream tasks.

The BERT was fine-tuned, as illustrated in Fig. 6. The commentary sentence was fed into the model to obtain the vector for each word. Moreover, a linear layer was added for classification. Finally, in the case of α_1 exceeding α_2 , the sentence was selected. Otherwise, it was not selected.

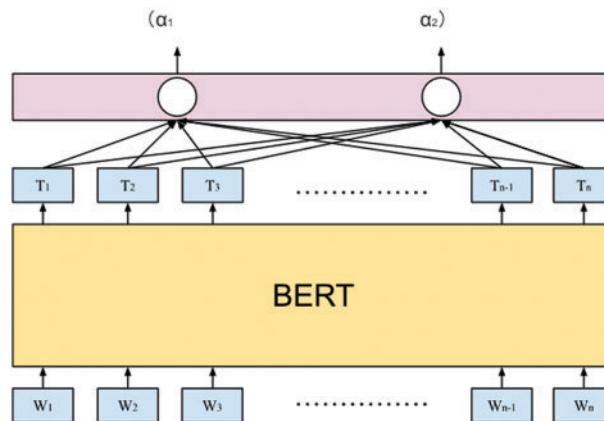


Figure 6: Classification model. Commentary sentences were fed into BERT, and a linear layer was added at the end of the model to facilitate classification

3.5 Rewriter

Rewriter was used to bridge the gap between commentary' writing styles and human-written news. In this section, the rewriter is similar to the polish service, and the generated sentences still have a high degree of overlap with the input text. Therefore, a text-editing model was implemented into the rewriter.

3.5.1 Task Formulation

In the previous process, the set of sentence pairs was obtained $\text{DataSet}_{\text{for_rewriter}} = \{(c_1, n_1), \dots, (c_i, n_i)\}$ where (c_i, n_i) represent the i -th commentary sentence c and news sentence n . The rewriter should transform c into a sentence with a more similar writing style to n while keeping the meaning accurate.

3.5.2 Text-Editing Model

The LaserTagger was adopted to perform the task of text editing [21], and RoBERTa-tiny-clue was used as the pre-trained language model [22].

As illustrated in Fig. 7, the LaserTagger⁴ turns the generation task into a sequence labeling task. The approach provides two basic labels: *KEEP*, *DELETE* and a particular label *ADD* for generation. The steps are as follows: (1) Constructing a limited set of phrases. The words generated are all from this set. Therefore, the best case is that the most labeled data can be reconstructed with the least words. The solution proposed by LaserTagger is first used to find the longest common subsequence (LCS) between sentence c and sentence n . After that, the words in n that are not in the LCS were composed into a candidate word set. Finally, the words with top 85% frequency were selected to form a limited set of phrases. (2) The words contained in c and n were labeled as *KEEP*; otherwise, they were labeled as *DELETE*. In addition, *ADD* was added after the basic label where it was generated and combined into *KEEP-ADD* or *DELETE-ADD*. (3) Generate output based on label sequences.

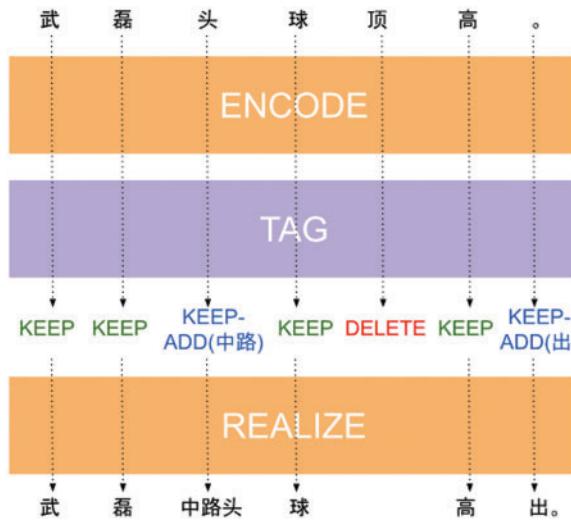


Figure 7: LaserTagger process. Commentary sentence was fed into the model to encode and marked with different labels. Thereafter, the sentence acted differently depending on the label

⁴LaserTagger refers to https://github.com/tongchangD/text_data_enhancement_with_LaserTagger

4 Experiments

In this section, multiple experiments are conducted to verify the effect of the selector andrewriter on the model's performance separately. First, the dataset is introduced in Section 4.1, and the metrics are introduced in Section 4.2. Finally, the automatic annotation experiment and the news generation experiment are discussed in Sections 4.3 and 4.4, respectively.

4.1 Dataset

This study conducts experimental investigations on the K-SportsSum dataset, which includes commentary and corresponding sports news releases of 7854 games [9]. Corrupted samples were removed during the data preprocessing to obtain high-quality training and testing datasets with 5540 and 340 games, respectively. In addition, since the public sports news datasets are all expanded based on previous studies, only the latest dataset was used in the experiments.

4.2 Metrics

Two evaluation metrics are used in our experiments, namely accuracy and ROUGE [23].

Accuracy is used to evaluate the number of correct automatic annotation results and defined as Eq. (2). s represents the number of samples, and function f is applied to evaluate the consistency between the automatic and manual annotation.

$$\text{Accuracy} = \frac{\sum_{i=1}^s f(\text{res}_{\text{automatic}}, \text{res}_{\text{manual}})}{s} \times 100\% \quad (2)$$

$$f(\text{res}_{\text{automatic}}, \text{res}_{\text{manual}}) = \begin{cases} 1, & \text{res}_{\text{automatic}} = \text{res}_{\text{manual}} \\ 0, & \text{others} \end{cases} \quad (3)$$

ROUGE evaluates the news generation effect by calculating the co-occurrence rate of n-grams. In our task, ROUGE-1, ROUGE-2, and ROUGE-L scores are used in our experiments. ROUGE-1 and ROUGE-2 calculate the recall rate of 1-gram and 2-gram between referenced news and generated news respectively. ROUGE-L calculates the overlap rate of the LCS.

4.3 Automatic Annotation Experiment

In this study, 40 samples were randomly picked from the dataset, and 2913 different sentences were obtained after sentence mapping. These sentences were annotated by humans as a reference to validate the automatic annotation results.

4.3.1 Baselines

The BERT-Score [24] and SBERT were selected as comparison baselines in this section. Generally, BERT-Score was adopted in the latest models [8], whereas the SBERT result was used as an ablation study to demonstrate the excellent performance of the proposed weight adjustment mechanism.

4.3.2 Analysis

The experimental results in Table 2 indicate that the SBERT + KEYWORDS achieves the best performance. Specifically, proposed algorithm improves the accuracy by 8% compared to the method which just considering similarity to annotation. This is because we added keywords match scores as rewards when annotating sentences. Moreover, compared to BERT-score, SBERT improves efficiency by about 15 times.

Table 2: Comparison of automatic annotation models. The table shows the time consumed to annotate 4030 sentences because some sentences are annotated multiple times. The accuracy is based on deduplicated sentences

Model	Accuracy	Time (s)
BERT-Score	90.68%	107
SBERT	93.92%	7
SBERT and KEYWORDS (Our)	98.90%	7

4.4 News Generation Experiment

The news generation experiment was designed to test the accuracy and efficiency of the rewriter. The SNG-TE was compared to the extractive, abstractive, and latest proposed models. Experimental results demonstrate that the proposed model performs better than others.

4.4.1 Baselines

Several generation models are selected for comparison. Extractive models include TextRank [25] and learning to rank (LTR) [3]. Generative models include long-short-term memory (LSTM) [26] and bidirectional and auto-regressive transformers (BART) [27]. In addition, the latest proposed models in sports news generation are compared, namely SportsSum [8] and knowledge enhanced summarizer (KES) [9].

4.4.2 Analysis

From the evaluation results in Table 3, ROUGE-1 score of SNG-TE was computed as 49.66, and the ROUGE-2 score was 21.47, both of which are better than the latest model. The proposed model retains the information available in the input and just needs to generate a small amount of content. Therefore, higher quality news can be generated.

Table 3: News generation results. The proposed model is compared with extractive models, abstractive models, and the latest proposed models

Method	Model	ROUGE-1	ROUGE-2	ROUGE-L
Extractive models	TextRank	20.57	6.01	20.23
	LTR	23.84	6.177	22.69
Abstractive models	LSTM	30.17	10.23	28.97
	BART	41.54	17.91	32.08
Two-step generation framework	SportsSum	44.89	19.04	44.16
	KES	48.79	21.04	47.17
	SNG-TE (Our)	49.66	21.47	35.18

It was also noticed that the ROUGE-L score is lower than the recently proposed models. This is because some specific terms were used to replace the expression of information in human-written sports news. For example, "第 92 分钟" could be replaced by "伤停补时", and "大禁区" could be

replaced by "20 码处", etc. However, the SNG-TE model can only select words from a limited set of phrases which does not include these terms, so that unable to express information in specific words. Although SNG-TE obtained the ROUGE-L score of only 19.99, the generated news can still accurately express the sports game, as shown in [Table 4](#).

Table 4: ROUGE-L analysis example. Although the ROUGE-L score is only 19.99, the generated sentences still accurately express the sports games

	Sample
Live commentary	第 92 分钟，巴塞罗那球员梅西大禁区外尝试左脚射门,可惜皮球偏出球门。
Model generation	第 92 分钟，梅西大禁区外左脚再射，可惜皮球偏出。
Manual writing	伤停补时，梅西 20 码处任意球直接射门高出。
ROUGE-L	19.99

Furthermore, the time consumption of SNG-TE and KES was recorded [9] to generate the 340 sports news in the testing set. As illustrated in [Fig. 8](#), the SNG-TE model took 452 s, and KES took 6819 s, indicating that the efficiency of the proposed model is better. The proposed model can retain the same words in the original and generated sentences without regenerating every word, improving efficiency by about 15 times.

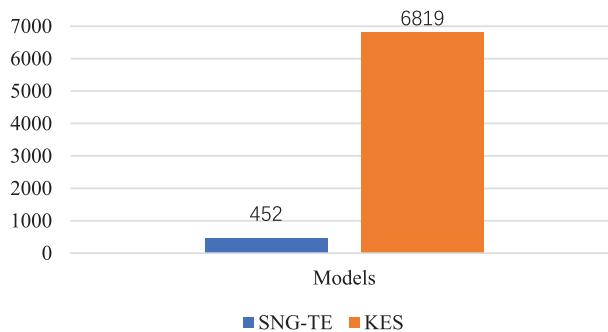


Figure 8: Efficiency of sports news generation. The recorded time consumption of SNG-TE and KES was used to generate the same amount of sports news. It demonstrates that the proposed model saves nearly 15 times as much time

5 Discussion

The main limitations in the automatic generation of sports news are the inability to select suitable sentences and the poor quality of the results. The proposed model provides solutions to the above two limitations.

Previous studies just considered the similarity to select suitable sentences. The SNG-TE model also considered keywords contained when selecting the sentences. [Table 2](#) illustrates that the accuracy of the proposed algorithm is 5% higher than the method that only considers the similarity.

Previous studies have shown that hallucinations cause the poor quality of generated news while using the Seq2Seq model. This study found that many of the sentences in the two paragraphs were

highly overlapping. Therefore, the text-editing model was introduced to generate sports news. As shown in [Table 3](#), the text-editing model achieves improved performance.

On the other hand, the proposed model also has some shortcomings. As shown in [Table 4](#), human writing uses specific expressions to increase the readability of news, which cannot be detected by the model. In addition, as shown in [Fig. 8](#), although LaserTagger saves time, it still generates text through autoregressive mode and cannot perform parallel inference acceleration.

6 Conclusion

This study proposes an approach to sports news generation that includes a selector and a rewriter. As part of the study, the commentary sentences in the selector were automatically annotated by the algorithm. Additionally, the BERT was fine-tuned for sentence classification. The text-editing model in the rewriter was applied to keep the generation process controllable while retaining valuable information in the input. In general, the study results have shown that in the experiment of automatic annotation, the algorithm's accuracy was increased by about 8% compared to the previous methods. On the other hand, in the news generation experiment, the ROUGE-1 score was computed as 49.66, and the ROUGE-2 score was 21.47, both of which are better than the latest model. Moreover, the proposed model reduced the required duration by nearly 15 times. Accordingly, the experimental results demonstrate that the proposed model improves the hit rate of important sentences and the efficiency of sports news generation while ensuring the accuracy of sports games. Further work is still needed to utilize the long-text pre-training model in solving the task of news generation and extend it to other fields.

Funding Statement: The works funded by the Research Project of Natural Science at Anhui Universities in 2021, Research on relation extraction of emergency plan knowledge graph based on deep embedding clustering (No. KJ2021A0994).

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] R. Tang, K. Zhang, S. Na, M. Yang, H. Zhou *et al.*, “Football news generation from Chinese live webcast script,” in *Natural Language Understanding and Intelligent Applications*, Kunming, China, pp. 778–789, 2016.
- [2] J. Yao, J. Zhang, X. Wan and J. Xiao, “Content selection for real-time sports news construction from commentary texts,” in *Proc. of the 10th Int. Conf. on Natural Language Generation*, Santiago de Compostela, Spain, pp. 31–40, 2017.
- [3] J. Zhang, J. Yao and X. Wan, “Towards constructing sports news from live text commentary,” in *Proc. of the 54th Annual Meeting of the Association for Computational Linguistics*, Berlin, Germany, pp. 1361–1371, 2016.
- [4] K. Zhang, J. Zhou, W. Wang, X. Lv, W. Zhang *et al.*, “Research on automatic writing of football news based on deep learning,” in *2017 2nd Int. Conf. on Computer Engineering Information Science and Internet Technology*, Sofia, Bulgaria, pp. 233–242, 2017.
- [5] L. Zhu, W. Wang, Y. Chen, X. Lv and J. Zhou, “Research on summary sentences extraction oriented to live sports text,” in *Natural Language Understanding and Intelligent Applications*, Kunming, China, pp. 798–807, 2016.

- [6] M. Liu, Q. Qi, H. Hu and H. Ren, “Sports news generation from live webcast scripts based on rules and templates,” in *Natural Language Understanding and Intelligent Applications*, Kunming, China, pp. 876–884, 2016.
- [7] X. Lv, X. You, W. Wang and J. Zhou, “Generate football news from live webcast scripts based on character-CNN with five strokes,” *Journal of Computers*, vol. 31, no. 1, pp. 232–241, 2020.
- [8] K. Huang, C. Li and K. Chang, “Generating sports news from live commentary: A Chinese dataset for sports game summarization,” in *Proc. of the 1st Conf. of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th Int. Joint Conf. on Natural Language Processing*, Seattle, United States, pp. 609–615, 2020.
- [9] J. Wang, Z. Li, T. Zhang, D. Zheng, J. Qu *et al.*, “Knowledge enhanced sports game summarization,” in *Proc. of the Fifteenth ACM Int. Conf. on Web Search and Data Mining*, New York, NY, United States, pp. 1045–1053, 2022.
- [10] Y. Keneshloo, T. Shi, N. Ramakrishnan and C. K. Reddy, “Deep reinforcement learning for sequence-to-sequence models,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, no. 7, pp. 2469–2489, 2019.
- [11] S. K. Kiran, M. Shashi and K. Madhuri, “Multi-stage transfer learning for fake news detection using AWD-LSTM network,” *International Journal of Information Technology and Computer Science (IJITCS)*, vol. 14, no. 5, pp. 58–69, 2022.
- [12] A. Zakharchenko, T. Peráček, S. Fedushko, Y. Syerov and O. Trach, “When fact-checking and ‘BBC standards’ are helpless: ‘Fake newsworthy event’ manipulation and the reaction of the ‘high-quality media’ on it,” *Sustainability*, vol. 13, no. 2, pp. 573, 2021.
- [13] K. M. Alzhrani, “Political ideology detection of news articles using deep neural networks,” *Intelligent Automation & Soft Computing*, vol. 33, no. 1, pp. 483–500, 2022.
- [14] A. A. Hnaif, E. Kanan and T. Kanan, “Sentiment analysis for arabic social media news polarity,” *Intelligent Automation & Soft Computing*, vol. 28, no. 1, pp. 107–119, 2021.
- [15] C. Oh, J. Choi, S. Lee, S. Park, D. Kim *et al.*, “Understanding user perception of automated news generation system,” in *Proc. of the 2020 CHI Conf. on Human Factors in Computing Systems*, New York, United States, pp. 1–13, 2020.
- [16] J. Wang, Z. Li, Q. Yang, J. Qu, Z. Chen *et al.*, “Sportssum2.0: Generating high-quality sports news from live text commentary,” in *Proc. of the 30th ACM Int. Conf. on Information & Knowledge Management*, Gold Coast, Queensland, Australia, pp. 3463–3467, 2021.
- [17] J. Wang, T. Zhang and H. Shi, “GOAL: Towards benchmarking few-shot sports game summarization,” arXiv preprint arXiv:2207.08635, 2022.
- [18] N. Reimers and I. Gurevych, “Sentence-bert: Sentence embeddings using siamese bert-networks,” in *Proc. of the 2019 Conf. on Empirical Methods in Natural Language Processing*, HongKong, China, pp. 671–688, 2019.
- [19] J. Devlin, M. Chang, K. Lee and K. Toutanova, “Bert: Pre-training of deep bidirectional transformers for language understanding,” in *Proc. of NAACL-HLT*, Minneapolis, United States, pp. 4171–4186, 2018.
- [20] L. Yao, C. Mao and Y. Luo, “KG-BERT: Bert for knowledge graph completion,” arXiv preprint arXiv:1909.03193, 2019.
- [21] E. Malmi, S. Krause, S. Rothe, D. Mirylenka and A. Severyn, “Encode, tag, realize: High-precision text editing,” in *Proc. of the 2019 Conf. on Empirical Methods in Natural Language Processing and the 9th Int. Joint Conf. on Natural Language Processing*, HongKong, China, pp. 5054–5065, 2019.
- [22] L. Xu, X. Zhang and Q. Dong, “CLUECorpus2020: A large-scale Chinese corpus for pre-training language model,” arXiv preprint arXiv:2003.01355, 2020.
- [23] C. Lin, “Rouge: A package for automatic evaluation of summaries,” in *42nd Annual Meeting of the Association for Computational Linguistics*, Barcelona Spain, pp. 74–81, 2004.
- [24] T. Zhang, V. Kishore, F. Wu, K. Weinberger and Y. Artzi, “Bertscore: Evaluating text generation with bert,” in *Int. Conf. on Learning Representations*, New Orleans, United States, 2019.

- [25] R. Mihalcea and P. Tarau, “Textrank: Bringing order into text,” in *Proc. of the 2004 Conf. on Empirical Methods in Natural Language Processing*, Barcelona, Spain, pp. 404–411, 2004.
- [26] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” in *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [27] Y. Shao, Z. Geng, Y. Liu, J. Dai, H. Yan *et al.*, “Cpt: A pre-trained unbalanced transformer for both Chinese language understanding and generation,” arXiv preprint arXiv:2109.05729, 2021.