Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 2374590, 8 pages
https://doi.org/10.1155/2022/2374590

Research Article

Automatic Generation of News Commentary on Social Media Based on Self-Attention Mechanism

Miao Qu

School of Journalism and Communication, Heilongjiang University, Heilongjiang 150000, China

Correspondence should be addressed to Miao Qu; 201811110911113@stu.hubu.edu.cn

Received 24 June 2022; Revised 29 July 2022; Accepted 8 August 2022; Published 16 September 2022

Academic Editor: Kuruva Lakshmanna

With development for social media, the technology of automatic comment generation for social media news is expected to generate great social and commercial value; so, this technology has attracted more and more attention from industry and academia. The automatic generation of social media news comments refers to the use of current popular technologies such as deep learning, NPL, and data mining to build high-performance algorithms to give machines the same ability to understand and express language as humans and to imitate human language habits to comment on social media news. Automatic generation of social media news comments is a challenging and understudied task in natural language processing. This work addresses the problem that social media news texts are too long and have complex contextual information, and existing work does not model the content structure of news well and proposes a topic encoding model. First, topic extraction and construction of news texts are carried out using the Latent Dirichlet Allocation (LDA) topic model to convert unstructured news texts into topic-sentence pairs. Second, a topic encoder is designed based on the self-attention mechanism, and topic-sentence pairs are input into the topic encoder. It goes through the sequence embedding module, the sentence encoding module, and the topic encoding module in turn to obtain the topic hidden state sequence. Finally, this paper designs a topic attention mechanism to generate diverse comments by giving different attention to each topic in the decoding part. Comprehensive and systematic experiments verify the effectiveness of this work and improve the performance of automatic generation of social media news comments.

1. Introduction

Information, ideas, career interests, and other kinds of expression are all made possible via the Internet and virtual communities, thanks to social media, an interactive computer-mediated technology. In this age of new media, social media has progressively risen to prominence as the primary means through which people gather and share information and express their views, thanks to the medium's inherent interactivity, immediacy, and accessibility. Compared with traditional counterparts, social media has broken through its functional boundaries and has become the main channel for people to obtain information. Social media connects all areas of social life through the Internet to infinitely connect and optimize our daily lives and work. Contemporary trends in global social media and platforms are characterized by their diverse, multiquantified, and multifaceted production. Social media and platforms are developing towards multitasking, multifunctional, mobile terminal, and highly open. Every day, a large number of active users and data streams converge on these media and platforms.

News sites set up a comment function, hoping to expand their news content and promote their audience engagement by encouraging users to browse comments, share information, and discuss news with each other. With the increasing popularity of news comments on the Internet, building an automatic generation system for news comments has become a research hotspot. A text generation system is a technology that generates readable textual representations by taking received language, data, images, and other forms of information as raw input. It is an important research branch in natural language processing. The news information on the Internet has a large audience, fast speed, strong timeliness, and rich news content, which brings a large number of news comments. These comments provide an expandable perspective that greatly facilitates audience engagement. Online news reviews allow users to express their opinions under each news article, demonstrate their attitudes, and communicate with other audiences. These open conversations encourage readers to exchange ideas, share new information, engage others, and encourage them to browse the page. Automatically generating news comments is one of research fields concerned with NLP, which has established an effective channel for news event influence analysis and public opinion evaluation.

The traditional automatic generation of news comments is mainly based on statistical machine learning for feature engineering, which is realized by means of retrieval. That is, given a piece of news, use the similarity calculation method to retrieve some news related to it from the news corpus. Then, from the comments corresponding to the retrieved news, the most relevant comment to the given news is selected as the generated comment. Although the quality of the generated comment is guaranteed by retrieval because it is a real comment, this retrieval-based automatic comment generation method is largely dependent on the type of news corpus. When the given news is very different from the retrieved news corpus, it is likely that no suitable reviews will be found; so, this method is not very scalable. In the context of the rise of deep learning, researchers try to generate news reviews based on deep learning. This method is not limited by news corpus and can generate diverse comments flexibly. However, building a news comment automatic generation system based on deep learning is a relatively new task in the field of NLG. Compared with the generation tasks like machine translation and dialogue generation, the research is still insufficient. Most of the current deep learning-based NLG methods use the Seq2Seq model. The task of automatic news comment generation involves a variety of cognitive abilities, such as understanding the meaning of the article, proposing a specific point of view, and expressing natural language; so, this task poses new challenges to the current Seq2Seq model. The task of automatic news comment generation is different from the summary generation task, and the comments it generates must cover all the important ideas of the article. It is also different from the product review task, where the input of product reviews is structured data, while the input of the automatic news review task is plain text. Plain text is vastly different in both structure and information than structured data; so, this constitutes a larger input space for exploration. The current challenges in the automatic generation of news comments are as follows. First, news texts are too long and contain too much information, and it is difficult for existing methods to capture key information in news. Conversely, while news headlines are a very important resource of information, using news headlines alone may be too short to provide enough information. Second, news articles and comments are not a pair of parallel texts, and news articles are much longer than comments and contain a lot of information unrelated to comments. Therefore, it is not suitable to directly apply the Seq2Seq model, which has been shown to be effective on other NLG tasks, to the task of automatic news comment generation. This may generate a lot of noise, resulting in insufficient sentence fluency. Third, users pay attention to different aspects of the news or have some emotional tendencies when posting comments, which makes the content of comments very diverse.

2. Related Work

NLG is a subtopic of artificial intelligence as well as computational linguistics. NLG aims to generate understandable text in human language. The development of NLG will help to build powerful intelligent systems can understand and synthesize language [1]. RNNs have shown good performance in text generation, an advantage that has led more and more researchers to explore various NLG tasks. In [2], authors used RNN to study product-specific review text generation. It utilizes a character-level RNN to generate relevant and coherent text given auxiliary information such as sentiment or topic. It uses a simple input replication strategy that preserves the signal of the auxiliary input over a wider sequence interval than backpropagation training through time. In particular, by contrast, its model can directly output categories and sentiments and corresponding review ratings. Authors have adopted [3] method to generate reviews based on user ratings in different aspects. In [4], authors used RNN to study the possibility of automatically and quickly generating hundreds of fake restaurant reviews. It generates reviews based on expected ratings and restaurant categories, a key feature of which is experimental evaluation, which involves human users. They judge the discriminative ability of users by showing them a series of reviews including real reviews and machine-generated reviews. Users are unaware that there are machine-generated reviews, and that they are participating in the assessment. Based on user identity, product identification, and user-specified rating, authors presented a system [5] to create an automatic review generator utilizing LSTM to generate tailored reviews. It is even better when you combine it with a sentiment analysis system that assigns ratings to reviews based on what they say. An additional loss term has been included to ensure that the sentiment ratings of generated reviews are round-robin consistent with conditional ratings used to generate them. The encoder-decoder in [6] expands small phrases that are entered into the system to generate customized reviews. The aspect encoder, which learns user and item representations, also includes information at the aspect level. The outputs of many encoders are monitored by an attention fusion layer, which is used in the control generation process.

In [7], authors proposed a framework that encodes the context into a vector. During decoding, contextual information is processed through a gating mechanism, which resolves the long-range dependency problem caused by lengthy sequences. In [8], authors have proposed a model for generating product reviews based on attribute information such as user, product, and rating which were proposed. The sequence decoder generates reviews by conditioning the output on these vectors, which the attribute encoder uses to learn how to represent input attributes as vectors. To ensure that the output and input properties are consistent, an attention mechanism has been included. In [9], authors projected a task of quantifiable sequence editing: edit an input sequence to generate an output sequence satisfying a given numerical value, measure a certain property of the sequence, and ask to preserve the main content of the input sequence. Authors have proposed a user preference-aware review generation model [10], which uses user preference words to improve the diversity of generated reviews. When decoding, it not only generates words according to the context vector but also generates words according to the user's preference. By considering the user's preference words, a personalized comment with the user's preference is generated. In [11], authors projected a method for combining retrieval model and generative model in view of the low quality of existing training data. It believes that the more likes a comment has, the better the quality of the comment. Therefore, the author first uses the likes of users to score the quality of the comment and proposes a scoring model for the comment. In [12], authors stated that in practical situations, online news usually contains multiple thematic content. Graphic news, for example, contains a large number of images in addition to text, and it believes that content other than text is also important because they are not only more attractive to readers but also provide key information. Therefore, a coattention model is proposed to capture the dependencies between textual and visual information, aiming to publish reviews by integrating multiple topical content. In [13], authors used the existing Seq2Seq model to construct four emotional language models and builds a system to generate emotional comments from newspaper articles. In [14], authors proposed a personalized comment generation network based on the real comments of users on Weibo and the corresponding user profiles. It utilizes user functions embedded in gated memory to personalize modeling based on user information to generate user-relevant reviews. In [15–17], authors proposed a model, which introduced topic attention mechanism algorithm and syntactic structure information to generate topic-specific Chinese comments.

3. Method

This work proposes a topic encoding model. First, topic extraction and construction of news texts are carried out using the LDA topic model to convert unstructured news texts into topic-sentence pairs. Second, a topic encoder is designed based on the self-attention mechanism, and topic-sentence pairs are input into the topic encoder. It goes through the sequence embedding module, the sentence encoding module, and the topic encoding module in turn to obtain the topic hidden state sequence. Finally, this paper designs a topic attention mechanism to generate diverse comments by giving different attention to each topic in the decoding part.

3.1. LDA Model. The LDA model is an unsupervised learning model that contains text topic expression functions that have been studied more in recent years. Its essence is to use the distribution rules of words in the text to achieve clustering processing of words with similar distribution. A generative statistical model called Latent Dirichlet Allocation (LDA) describes a collection of observations by means of unobserved groups, and each specific group describes how certain aspects of the data are similar. It can be defined as a probabilistic topic model and can find hidden and difficult to find topic patterns in all discrete data. From Figure 1, it can be seen that the LDA model is a three-layer Bayesian model based on document-topic-word. Its central idea is that each document can be reflected by a composite distribution of several implicit topics, and all topics are found through the probability distribution of all words in the word set. LDA has two priors, one is based on text and topic, and the other is based on topic and words. The prior distribution function is

$$\mathrm{Dir} = \frac{\Gamma\left(\sum_i \alpha_i\right)}{\Pi_i \Gamma(\alpha_i)} \prod_{i=1}^{T} \theta_i^{\alpha_i - 1}. \tag{1}$$

As shown in Figure 2, LDA is a very representative directed probabilistic graphical model. α expresses the relative strength and weakness of all implicit topics in the corpus. β represents the probability distribution of all hidden topics itself, the random variable θ represents the document layer, and its size reflects the share of each hidden topic in the target text. In the word layer, Z is used to reflect the ratio of the hidden topics that the target document is divided into each word, and the expression form represented by W belongs to the target text.

In most cases, the characteristics of the LDA model in the same topic will have a great correlation, such as the characteristics of high coexistence and high correlation. The LDA topic model's most important benefit is its ability to effectively reduce the dimensionality of words, turning an initial high-dimensional word space into a low-dimensional space with fewer topic components. On the one hand, this dimensionality reduction process can avoid the overcomplicated similarity calculation, and on the other hand, it can also solve the problem of data sparsity that often occurs in text models. The topic model can map all texts from high-dimensional space to low-dimensional space and convert them into vector form in the topic space. Low-dimensional topic spaces are usually dense, which can motivate exploration to discover potentially unobvious relationships. In this way, the interference caused by data sparsity to the text similarity measurement may be avoided to a certain extent, and the potential meaning in the text set may be more completely discovered. The LDA topic model is an extended development of probability theory, which uses probability and statistics theory based on the bag-of-words hypothesis to process text, reflecting that a text can be represented according to irregular combinations. Different from linguistic text mining algorithms based on grammatical rules, LDA's method of discovering implicit semantics is based on the frequency of occurrence of related words in the text. This approach will yield more efficient results in very grammatically irregular text sets. From another perspective, the LDA model uses a probability distribution algorithm to process text. One probability represents

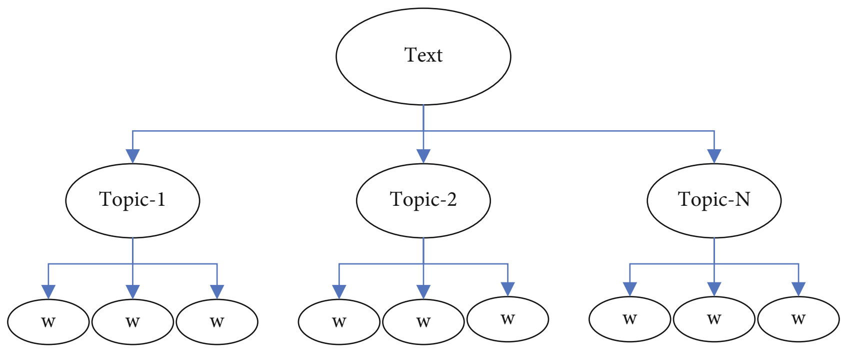

Figure 1: LDA implicit theme topology.

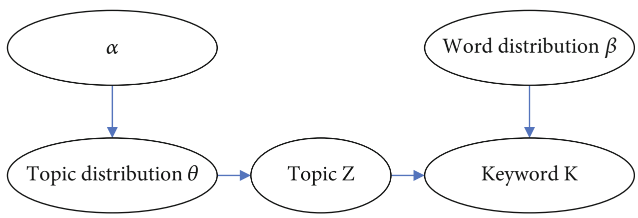

Figure 2: LDA graph model.

a topic, and the text is assigned to different topics due to different probabilities, which satisfies the probability distribution of real data. The algorithm of probability distribution can more reasonably reflect and concrete the uncertain factors already contained in the text and reduce the noise and interference.

3.2. Attention Mechanism. The Seq2Seq model encodes the input sequence into a semantic vector through the encoder, which is then decoded by the decoder. However, as the sequence length increases, the semantic information in the encoding process will be gradually weakened with the iteration of the sequence moment. The resulting semantic vector cannot represent the full information of the input, which leads to low quality of the generated sentences. For example, in the task of machine translation, when the sentence to be translated is too long, the fixed semantic vector cannot contain all the information, which reduces the translation accuracy. In this way, the attention mechanism assigns different attention to each word in the input English sequence at different times of decoding to align the source language and the target language, thereby generating a more fluent translation. Since then, the attention mechanism has been widely used in various NLP tasks based on neural networks such as RNN/CNN.

The attention mechanism which can be regarded as the calculation process of mapping query into key-value pairs. Specifically, the first is to calculate the weight of the query and each key through the similarity function. The weights are then normalized using the softmax function. Finally, the weight and value are weighted and summed to get the attention output.

There are many ways to implement the attention mechanism, and the different implementation ways are mainly reflected in the similarity algorithm between the query vector and the key vector. The content-based attention mechanism calculates the similarity score between the query vector and the key vector through the cosine similarity calculation.

$$\mathrm{Score}(K, Q) = \mathrm{cosine}(K, Q). \tag{2}$$

The three attention mechanisms of dot, general, and concat are

$$\mathrm{Score}(K, Q) = \begin{cases} K^T Q, & \text{dot}, \\ K^T W_a Q, & \text{general}, \\ v_a^T \tanh\left(W_a[K, Q]\right), & \text{concat}. \end{cases} \tag{3}$$

The dot attention mechanism is to directly multiply the K vector and the Q vector without any parameters in the calculation process; so, the memory occupied by the model is very small. The general attention mechanism establishes a linear relationship between the K vector and the Q vector by introducing parameters. The concat attention mechanism exploits the correlation between the K vector and the Q vector more fully by introducing two parameters and the tanh activation function. However, this will increase the number of parameters, and the memory occupied by the model will increase.

The self-attention mechanism is calculated as

$$\mathrm{Attention}(Q, K, V) = \mathrm{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right) V. \tag{4}$$

Do the dot product of Q and K first and then divide by a scale to avoid the dot product being too large. The results are then normalized using the softmax function and finally multiplied by V for weight summation. In order to extract the features comprehensively, through the multihead attention mechanism, K, Q, and V are obtained by a linear mapping to several K, Q, and V. Each head computes self-attention and combines them to get the final vector representation.

3.3. Topic Encoder Model. This section proposes a topic encoding model, the structure of which is shown in Figure 3, which consists of two parts: topic encoder and decoder with topic attention mechanism. Among them, in the encoder part, this chapter designs a topic encoder to organize news structure, analyze, and understand the

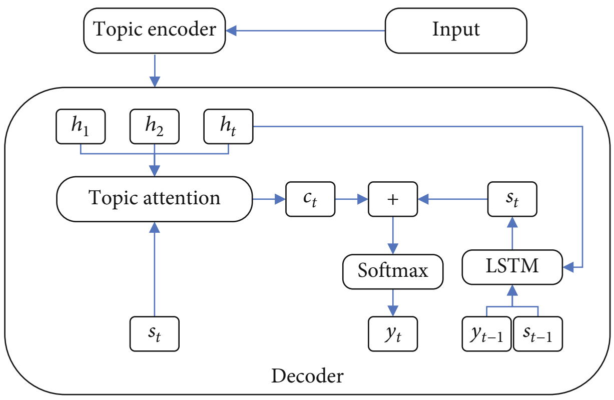

Figure 3: Topic encoder model.

relationship between news content and topics. The encoder, which comprises of three modules, introduces the location encoding method and the multihead self-attention mechanism in the transformer model: sequence embedding module, sentence encoding module, and topic encoding module. A unidirectional LSTM is used in the decoder part to generate reviews, among which, a topic attention mechanism is designed in the proposed method. The topic encoder is used by this model to extract the topic hidden state sequence, and it assigns varying degrees of priority to each subject and its associated sentences to provide a variety of insightful remarks.

Firstly, the topic extraction and construction of news text are carried out by using related algorithms, the unstructured news data is transformed into topic-sentence pairs, and then the topic-sentence pairs are input into the topic encoder for semantic feature extraction. The topic encoder sends the topic-sentence pair to the topic encoder, and the sequence embedding module calculates the word vector representation in turn. The sentence vector representation sequence is calculated by the sentence encoding module, and the topic hidden state sequence is calculated by the topic encoding module. At the same time, considering that news headlines are an important information resource, this chapter also considers that news headlines are input into the topic encoder, and the vector representation of news headlines is calculated by sequence embedding module and sentence encoding module in turn. Finally, in the decoding part, the news headline vector representation obtained by the topic encoder acts on the decoder to hide the initial state to strengthen the influence of important news information on comment generation. Through the topic attention mechanism, each topic and related sentences are given different degrees of attention at each time step of decoding, and the topic vector representation is calculated.

3.4. Topic Extraction and Construction. News contains multiple topic information, and it is difficult for existing models to distinguish each topic and the corresponding text content. Therefore, this work first extracts and constructs the topic of the news text and obtains a structured topic-sentence pair as the input of the topic encoder. The implementation process is shown in Figure 4.

3.5. Topic Encoder. The topic encoder proposed in this chapter consists of three modules, namely, sequence embedding module, sentence encoding module, and topic encoding module. The structure is shown in Figure 5. The input topic-sentence pairs of this module, and considering that news headlines are also an important information resource, the news headlines are also used as the input of the encoder. Among them, the sequence embedding module introduces the position encoding mechanism of the transformer model, and the position of each word is input into the model as a feature, which is conducive to better understanding the contextual relationship between words. The sentence encoding module and the topic encoding module introduce the multihead self-attention mechanism of the transformer model, because the self-attention mechanism in the transformer model can enable each word in the input to ignore the distance limitation and encode other words. The working process of the three modules are represented in the below equations, where the combination of embedding and position vector provides the value for v_j^i. Similarly, the self attention process is being calculated by means of the continuous vectors.

$$\begin{gathered} v_j^i = \text{Embedding}\left(s_j^i\right) + \text{Position}\left(s_j^i\right), \\ \left(h_1^i, h_2^i, \cdots, h_m^i\right) = \text{SelfAttention}\left(v_1^i, v_2^i, \cdots, v_m^i\right), \\ h_i^s = \text{MLP}\left(h_1^i, h_2^i, \cdots, h_m^i\right), \\ \left(h_1^T, h_2^T, \cdots, h_k^i\right) = \text{SelfAttention}\left(h_1^S, h_2^S, \cdots, h_k^S\right). \end{gathered} \tag{5}$$

This work also inputs the news headlines into the sequence embedding module and the sentence encoding module and obtains the news headline vector. The calculation process is consistent with the above. The resulting news headline vector is used as the initial vector for the first step of the decoder. Because from the actual news data, the content of news headlines is often the most attractive point of a piece of news.

3.6. Decoder. On the decoding side, the decoder of this model is implemented based on a two-layer unidirectional LSTM. At any moment of decoding, the decoder reads the hidden state at the previous moment and outputs the corresponding word vector at the previous moment to calculate the hidden state. This work uses the news headline vector obtained by the topic encoder as the initial state of the decoder hidden state.

In order to characterize the soft-aligned relationship between the output hidden state and its related topics, a topic attention mechanism is designed in this work. The purpose is to establish a weight relationship between the output and each topic at the decoding end after obtaining the topic hidden state sequence through the topic encoder and assign different weights to each topic hidden vector representation. This allows the model to know which topic is more closely related to the word to be output at the time

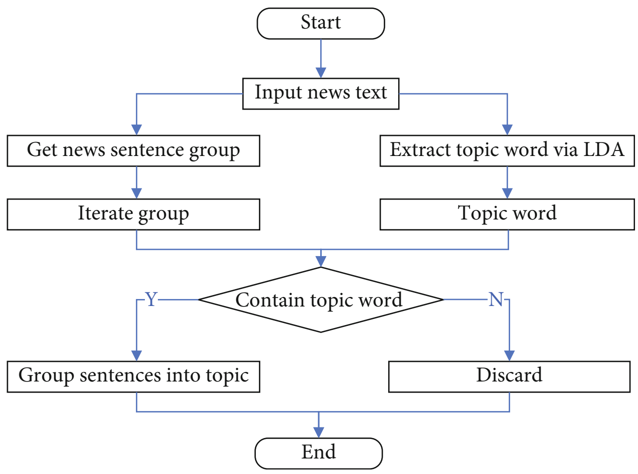

FIGURE 4: Topic extraction and construction pipeline.

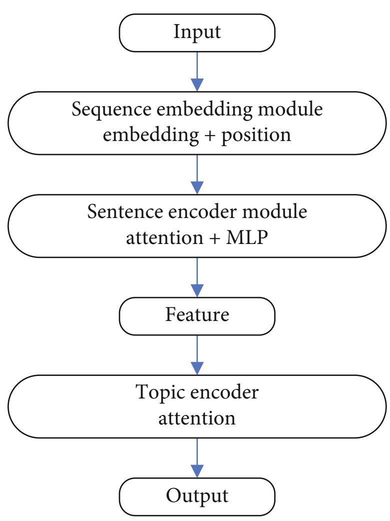

FIGURE 5: Topic encoder structure diagram.

of decoding, thereby increasing the diversity of comments.

$$
\begin{aligned}
e_{ti} &= v_a^T \tanh\left(W_a s_t + U_a h_i^T\right),\\
a_{ti} &= \operatorname{softmax}(e_{ti}),\\
c_t^T &= \sum_i a_{ti} h_i^T.
\end{aligned}
\tag{6}
$$

Finally, at the decoding end, the topic vector and the hidden state vector of the decoder are spliced together, and a new hidden state vector is obtained through a linear change. Then, go through a softmax layer to get the probability distribution of the output words.

4. Experiment

4.1. Dataset and Evaluation Metric. For the proposed work, the data set has been considered from the crawler technology from the network for network training and testing, where 90,863 samples are used for training and 30,518 samples are used for testing. The performance evaluation metrics used in this work are BLEU and METEOR.

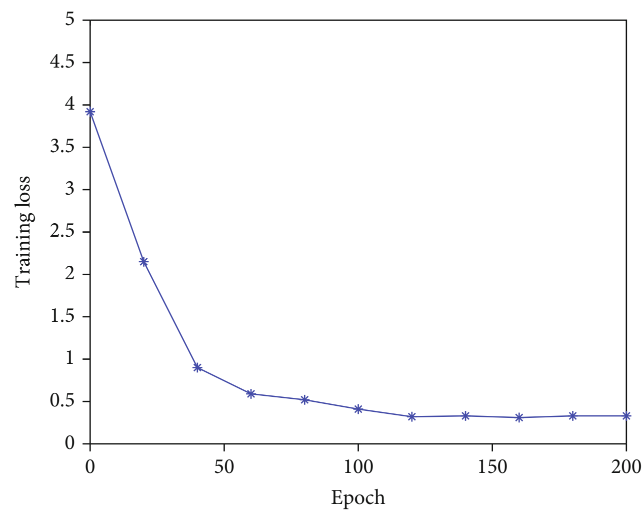

FIGURE 6: Network training analysis.

4.2. Network Training Analysis. First, the evaluation of the automatic comment generation model designed in this work is carried out to evaluate the network training. The loss in the network training process is demonstrated in Figure 6.

With the deepening of network training, the network loss gradually decreases. When the training epoch reaches around 120, the network converges.

4.3. Comparison with Other Methods. To verify the superiority of the topic encoding model proposed in this work for automatic generation of social media news comments, it is compared with other comment generation methods, and the comparison results are demonstrated in Table 1.

From the above table, it clearly shows that the proposed method is having good results and has achieved the highest BLEU and METEOR, and these results confirm that the proposed method has a better approach over the existing methods.

4.4. LDA Module Analysis. This work uses the LDA model to extract the subject words from the news topic. To verify the superiority of the LDA strategy, the performances without and with LDA are compared, respectively, and the results are demonstrated in Figure 7.

Compared with the data in the figure, when the LDA module is used, the indicators of BLEU and METEOR can be improved, respectively, which proves the correctness of using LDA.

4.5. Self-Attention Analysis. This work uses the self-attention mechanism to extract the corresponding text features. To verify its superiority over traditional attention, the performance of traditional attention and self-attention is compared, respectively. The comparison data is demonstrated in Figure 8.

Compared with the data in the figure, when self-attention is used, the indicators of BLEU and METEOR

Table 1: Comparison for different methods.

Method	BLEU	METERO
Attention	0.172	0.081
GANN	0.185	0.072
Grape2Seq	0.191	0.085
Proposed method	0.218	0.103

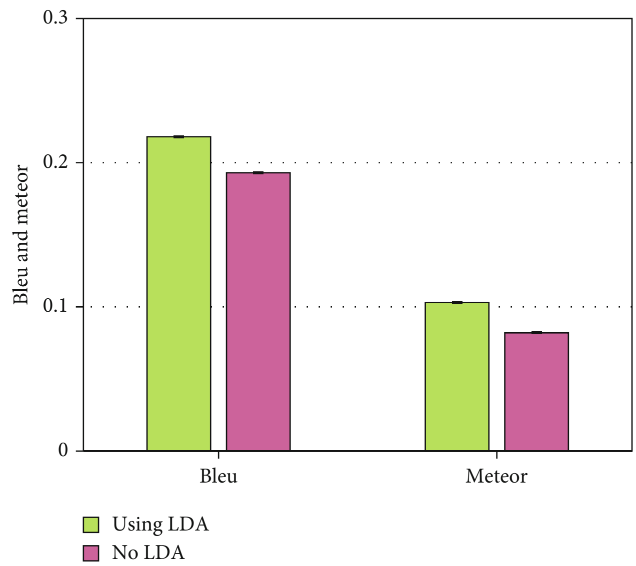

Figure 7: LDA module analysis.

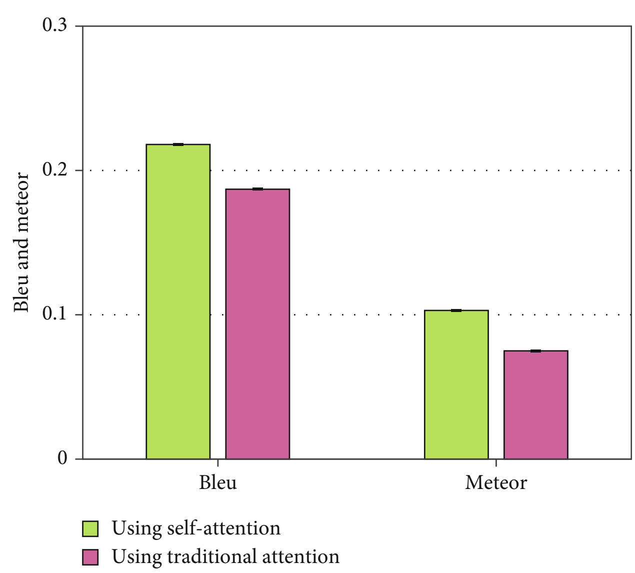

Figure 8: Self-attention analysis.

can be improved, respectively, which proves the correctness of using self-attention rather than traditional attention.

4.6. Topic Attention Analysis. This work uses the topic attention mechanism on the decoding side. To verify the superiority of this measure, we compare BLEU and METEOR without and using topic attention, respectively, as demonstrated in Figure 9.

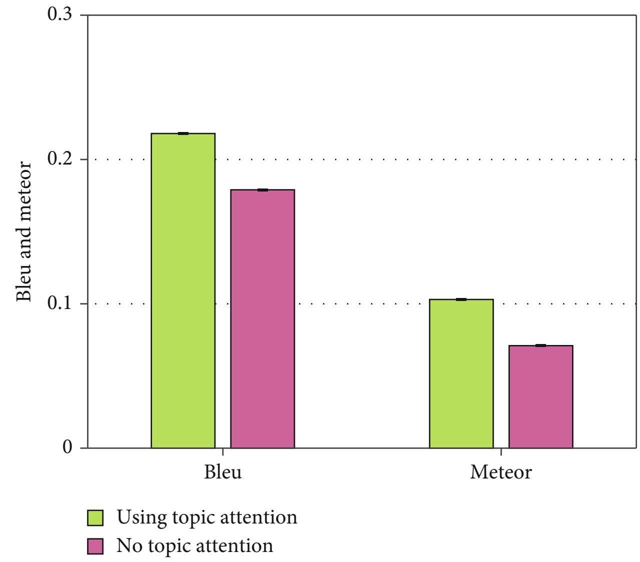

Figure 9: Topic attention analysis.

Table 2: Multihead attention analysis.

Head	3	4	5	6
BLEU	0.192	0.218	0.201	0.185
METEOR	0.893	0.103	0.095	0.869

Compared with the data in the figure, when topic attention module is used, the indicators of BLEU and METEOR can be improved, respectively, which proves the correctness of using this module.

4.7. Multihead Attention Analysis. This work uses a multihead self-attention mechanism to extract text features, in which the number of heads is variable. To verify the impact of different attention heads on performance, this work conducts comparative experiments as demonstrated in Table 2.

In the multihead self-attention mechanism, selecting different numbers of heads will also achieve different performance. And as the number of attention heads increases, the model performance first increases and then decreases, and the highest performance can be obtained when the value is 4.

5. Conclusion

Automatic generation of social media news comments is a challenging and understudied task in the field of natural language processing. With the explosion of information on the Internet, there is a growing need to build an automatic generation system for social media news comments. It helps increase user engagement and interactivity for news sites, while acting as a commenting assistant and enriching the chatbot's feature list. With the rise of deep learning, research on automatic generation of social media news comments has transitioned from retrieval-based models to sequence-to-sequence models based on deep neural networks. This work addresses the problem that social media news texts are too long and have complex contextual information, and the existing work does not model the content structure of news well and proposes a topic encoding model. First, topic extraction and construction of news texts are carried out using the LDA topic model to convert unstructured news texts into topic-sentence pairs. Second, a topic encoder is designed based on the self-attention mechanism, and topic-sentence pairs are input into the topic encoder. It goes through the sequence embedding module, the sentence encoding module and the topic encoding module in turn to obtain the topic hidden state sequence. Finally, this paper designs a topic attention mechanism to generate diverse comments by giving different attention to each topic in the decoding part. Comprehensive and systematic experiments verify the effectiveness of this work and improve the performance of automatic generation of social media news comments.

Data Availability

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

Conflicts of Interest

The author declares that he has no conflict of interest.

References

[1] L. W. Ku and H. H. Chen, "Mining opinions from the Web: beyond relevance retrieval," *Journal of the American Society for Information Science and Technology*, vol. 58, no. 12, pp. 1838–1850, 2007.

[2] Z. C. Lipton, S. Vikram, and J. McAuley, "Capturing meaning in product reviews with character-level generative text models," 2015, http://arxiv.org/abs/1511.03683.

[3] J. Hunter, Y. Freer, A. Gatt, E. Reiter, S. Sripada, and C. Sykes, "Automatic generation of natural language nursing shift summaries in neonatal intensive care: BT-Nurse," *Artificial Intelligence in Medicine*, vol. 56, no. 3, pp. 157–172, 2012.

[4] A. Berard, I. Calapodescu, M. Dymetman, C. Roux, J. L. Meunier, and V. Nikoulina, "Machine translation of restaurant reviews: new corpus for domain adaptation and robustness," 2019, http://arxiv.org/abs/1910.14589.

[5] A. Felfernig, G. Friedrich, D. Jannach, and M. Stumptner, "Consistency-based diagnosis of configuration knowledge bases," *Artificial Intelligence*, vol. 152, no. 2, pp. 213–234, 2004.

[6] H. Xu, H. Liu, W. Zhang, P. Jiao, and W. Wang, "Rating-boosted abstractive review summarization with neural personalized generation," *Knowledge-Based Systems*, vol. 218, article 106858, 2021.

[7] J. Tang, Y. Yang, S. Carton, M. Zhang, and Q. Mei, "Context-aware natural language generation with recurrent neural networks," 2016, http://arxiv.org/abs/1611.09900.

[8] A. Radford, R. Jozefowicz, and I. Sutskever, "Learning to generate reviews and discovering sentiment," 2017, http://arxiv.org/abs/1704.01444.

[9] Y. Liao, L. Bing, P. Li, S. Shi, W. Lam, and T. Zhang, "Quase: sequence editing under quantifiable guidance," in *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pp. 3855–3864, Brussels, Belgium, 2018.

[10] Y. Zhang, C. H. Zhang, and X. Shao, "User preference-aware navigation for mobile robot in domestic via defined virtual area," *Journal of Network and Computer Applications*, vol. 173, article 102885, 2021.

[11] S. I. Koshoridze, "New comment on "aqueous dispersions of nanobubbles: generation, properties and features" by A. Azevedo, R. Etchepare, S. Calgaroto, J. Rubio [Miner. Eng. 94 (2016) 29-37]," *Minerals Engineering*, vol. 142, no. 105877, p. 105877, 2019.

[12] T. Haije, B. O. Intelligentie, E. Gavves, and H. Heuer, "Automatic comment generation using a neural translation model," *Information and Software Technology*, vol. 55, no. 3, pp. 258–268, 2016.

[13] Y. Lin, P. Ren, Z. Chen, Z. Ren, J. Ma, and M. De Rijke, "Explainable outfit recommendation with joint outfit matching and comment generation," *IEEE Transactions on Knowledge and Data Engineering*, vol. 32, no. 8, pp. 1502–1516, 2020.

[14] G. Sridhara, *Automatic Generation of Descriptive Summary Comments for Methods in Object-Oriented Programs*, University of Delaware, 2012.

[15] W. Zeng, A. Abuduweili, L. Li, and P. Yang, "Automatic generation of personalized comment based on user profile," 2019, http://arxiv.org/abs/1907.10371.

[16] A. P. Rodrigues, R. Fernandes, A. Shetty, K. Lakshmanna, and R. M. Shafi, "Real-time twitter spam detection and sentiment analysis using machine learning and deep learning techniques," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 5211949, 14 pages, 2022.

[17] J. Pei, Z. Yu, J. Li, M. A. Jan, and K. Lakshmanna, "TKAGFL: a federated communication framework under data heterogeneity," *IEEE Transactions on Network Science and Engineering*, 2022.