

Research Article

Analysis of the Application of Feedback Filtering and Seq2Seq Model in English Grammar

Aizhen Zhang 

School of International Studies, Henan Finance University, Zhengzhou 450046, China

Correspondence should be addressed to Aizhen Zhang; zhangaz@hafu.edu.cn

Received 7 February 2022; Accepted 4 March 2022; Published 24 March 2022

Academic Editor: Kalidoss Rajakani

Copyright © 2022 Aizhen Zhang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Natural language processing (NLP) technology is widely used in grammatical error correction, but its error correction logic is complex and fault-tolerant, which leads to low accuracy. With the progress of deep learning and big data analysis technology, a new method is proposed in the technical means of English grammar error correction. This paper proposes a deep learning model-based feedback grammar error correction method, which can effectively improve the accuracy and tolerance of grammar error correction. Firstly, the **Seq2Seq** model with attention mechanism is proposed, and then the feedback filtering model is integrated, so that the existing errors or inefficient grammars can be corrected again, thus improving the efficiency of the model. Through a large number of text detection, the model proposed in this paper has high execution efficiency and application ability and can widely meet the needs of English translation and grammar correction.

1. Introduction

As an international language, English plays a vital role in international communication. With the economic globalization and the increasing frequency of international exchanges, the demand for talents who are proficient in English is gradually increasing, which makes many non-English-speaking countries pay more and more attention to English teaching and education. In China, English is a compulsory language, which runs through all stages of education in primary school, junior high school, senior high school, and university. However, due to the lack and uneven distribution of English teaching resources, the feedback period of English teaching effect is long, and the level of English learners cannot be improved rapidly. To solve this problem, relevant experts and scholars have applied computer technology to assist English learning and studied and designed many English-assisted learning systems. At present, English-assisted learning systems are mainly English grammar error correction systems, such as Literature [1, 2]. Based on English corpus and other resources and computer technology, through model training, English grammar can be quickly corrected, and the feedback cycle of English education and teaching

can be shortened. However, although these systems provide great convenience for English education and teaching, there are some problems such as low error correction accuracy and limited error correction ability. The reason is that most models based on English corpus resources depend on the quality of corpus resources, and the model training is complicated. Therefore, through the above analysis and research, combined with the characteristics and advantages of in-depth learning, an English grammar error correction and feedback method based on **Seq2Seq** model is proposed to improve the error correction ability and effect of the system and meet the actual needs of English-assisted learning.

2. Seq2Seq Model and Its Improvement

2.1. Basic Introduction of Seq2Seq Model. **Seq2Seq** model is a variant cyclic neural network for natural language processing, including encoder and decoder. Among them, the encoder is a cyclic neural network, which is mainly responsible for compressing the input sequence into vectors of specified length; the decoder is also usually a cyclic neural network, which is mainly responsible for generating the specified sequence [3, 4] according to the semantic vector.

The **Seq2Seq** model can learn the probability distribution and make predictions. The specific mathematical description is as follows.

If the input sequence is $x = |x_1, x_2, \dots|$ and the conversion rule is f , the hidden state HT at the current time can be expressed as

$$h_t = f(h_{t-1}, x_t), \quad (1)$$

where h_{t-1} represents the hiding state at the previous time.

According to the hidden state of each moment in Equation (1), all hidden states are summarized to obtain semantic vector C .

$$C = q(h_1, h_2, h_3, \dots, h_t). \quad (2)$$

The decoding stage can be understood as the inverse process of encoding the above codes. According to C and the generated output sequence y_0, y_1, \dots, y_{t-1} , the next output word y_t is predicted.

$$y_t = \arg \max p(y_t) = \prod_{t=1}^T p(y_t | \{y_1, y_2, \dots, y_{t-1}\}, C). \quad (3)$$

There are also calculation rules in the decoding process. h'_t is used to represent the hidden state in the process, including

$$h'_t = f(h'_{t-1}, y'_{t-1}, C). \quad (4)$$

2.2. Improvement of Seq2Seq Model. The research shows that although **Seq2Seq** model is similar to machine translation algorithm, its translation efficiency and error correction ability need to be further improved. Therefore, this study improves the model from three aspects: attention mechanism, addition layer standardization, and word embedding processing.

2.2.1. Attention Mechanism. Attention mechanism can be divided into local attention mechanism and global attention mechanism according to the calculation range of input sequence [5]. In this study, the problem of English grammar error correction needs to be dealt with according to the full text, so the global attention mechanism combined with **Seq2Seq** model is selected to deal with the problem of English grammar error correction. The specific operation process can be illustrated in Figure 1.

As can be seen from the figure, the value of the weight a_{ij} can be determined by the $i-1$ hidden state s_{i-1} and each hidden state variable in the input, and its calculation formula is as follows:

$$a_{ij} = \frac{\exp(e_{ij})}{\sum_{k=1}^{Tx} \exp(e_{ik})}, \quad (5)$$

$$e_{ij} = a(S_{i-1}, h_j). \quad (6)$$

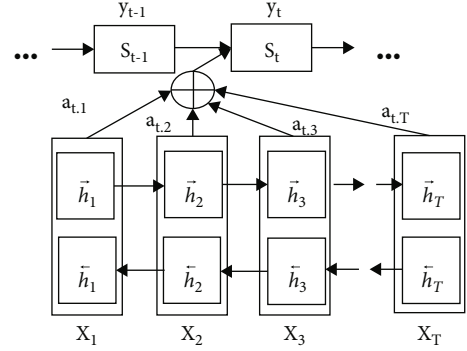


FIGURE 1: Global attention mechanism model.

2.2.2. Adding Layer Normalization. It is considered that the training efficiency of the model can be improved, and the training time of the model can be reduced by standardizing the input activation function of the model. In order to improve the training efficiency of **Seq2Seq** model, LN layer normalization is added to the model. LN layer normalization is a kind of horizontal normalization. By summing and variance operations on the inputs of the entire layer of neurons in the cyclic neural network in the model, and mapping the layer of neurons into the same distribution, the normalization terms shared by hidden elements are realized μ and σ , such as Equations (7) and (8). After layer normalization, the performance of cyclic neural network is better than that of original network.

$$\mu = \frac{1}{H} \sum_{i=1}^H x_i, \quad (7)$$

$$\sigma = \sqrt{\frac{1}{H} \sum_{i=1}^H (x_i - \mu)^2}. \quad (8)$$

2.2.3. Word Embedding Processing. Since the computer cannot directly process natural language, it is necessary to vectorize the input English text to convert it into data that can be directly processed by the computer [6]. In this study, Continuous Bag of Words Model (CBOW) is used to vectorize the English text of the input model. Its basic transformation idea is to calculate the probability of a word appearing according to the first N times or the last N words of a word, as shown in Figure 2.

3. Feedback Filtering Algorithm Model Based on N -Gram

Considering that the correction results of grammatical error correction model may not be correct or achieve ideal error correction effect, especially for grammatical errors of complex sentences, manual error correction ability is stronger. Therefore, in order to improve the error correction ability of the error correction model, we should consider accepting the modification suggestions submitted by users. However, due to the uneven English level of users, the model

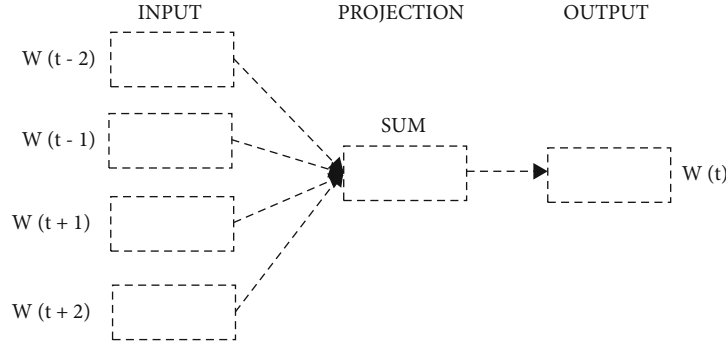


FIGURE 2: Structure diagram of CBOV model.

needs to screen the user's suggestions before accepting the user's modification suggestions. The simplest screening method is to compare which of the system and the user's modification suggestions is more credible, that is, to score the modified sentences. The sentences with higher scores are considered to be sentences with higher probability without language disorders, while the sentences with lower scores are not adopted. In this study, n -gram model, a basic language model used in the field of natural language processing, is selected to score sentences. The mathematical description of the model is as follows.

Suppose S is a string representing a statement; $P(s) = P(\omega_1, \omega_2, \dots, \omega_m)$ denotes the occurrence probability of a sentence composed of m words. First, N -gram model is used to model S . Then, according to the chain rule, such as Equation (9), the probability of S is calculated.

$$P(\omega_1, \omega_2, \dots, \omega_m) = P(\omega_1)P(\omega_2 | \omega_1)P(\omega_3 | \omega_2, \omega_1) \dots P(\omega_m | \omega_1, \dots, \omega_{m-1}). \quad (9)$$

According to the formula, the probability of the i -th word ω_i is related to the $\omega_1 \omega_2 \dots \omega_{i-1}$ word before it, and there is a dependency relationship between words, such as ω_i depends on ω_{i-1} before it. However, considering the large calculation of this formula, the calculation is difficult, and the parameters of n -gram model are difficult to obtain. Therefore, the study introduces Markov hypothesis, that is, the current word is only related to the previous N words, then, Equation (9) can be simplified to

$$P(\omega_i | \omega_1, \dots, \omega_{i-1}) = P(\omega_i | \omega_{i-n-1}, \dots, \omega_{i-1}). \quad (10)$$

Usually, $n = 1, 2, 3$. When $n = 2$, the calculation process is

$$P(\omega_1, \omega_2, \dots, \omega_m) = \prod_{i=1}^m P(\omega_i | \omega_{i-1}). \quad (11)$$

When $n = 3$, the calculation process is

$$P(\omega_1, \omega_2, \dots, \omega_m) = \prod_{i=1}^m P(\omega_i | \omega_{i-2} \omega_{i-1}). \quad (12)$$

Second, the maximum value of Equation (9) can be taken

according to the maximum likelihood estimation method. In the process of corpus training, the estimation can be made through the following equation:

$$P(\omega_i | \omega_{i-2} \omega_{i-1}) = \frac{c(\omega_{i-2} \omega_{i-1} \omega_i)}{c(\omega_{i-2} \omega_{i-1})}. \quad (13)$$

According to the trained n -gram model and the interface provided by SRILM, the sentence probability can be calculated simply.

Finally, by calculating the confusion degree of the sentence, the sentence can be scored. The calculation formula of confusion is as follows:

$$PP(S) = P(\omega_1, \omega_2, \dots, \omega_m)^{-1/m} = \sqrt[m]{\frac{1}{P(\omega_1, \omega_2, \dots, \omega_m)}}. \quad (14)$$

According to the chain rule, the above formula is rewritten to obtain:

$$PP(S) = \sqrt[m]{\prod_{i=1}^m P(\omega_i | \omega_{i-n-1}, \dots, \omega_{i-1})}. \quad (15)$$

As can be seen from the above formula, the smaller the degree of confusion, the greater the probability of sentences.

By comparing confusion degree and probability, it can be seen that confusion degree is more accurate in evaluating the model. The reason is that probability calculation is closely related to sentence length and word number, which affects the accuracy of model evaluation. Confusion degree can effectively reduce the influence of sentence length and word number on model evaluation, so its evaluation result is more accurate. Therefore, this study chooses confusion degree to judge the model.

4. Experimental Verification

4.1. Verification of English Error Correction Algorithm

4.1.1. *Data Set Source and Preprocessing.* NUCLE-release2.2 established by National University of Singapore is selected as the training corpus for the model verification data set of this study. The corpus includes 28 types of English grammatical errors and is manually labeled. Specific labeling [7,

8] examples are shown in Table 1, including sentence document ID, female tag, and sentence paragraph index ID. In addition, in order to facilitate the evaluation of the error correction results of this model, the data set was simply processed before the experiment. First, according to the annotation format of data set, it is transformed into m2 format to obtain the original English sentence and its grammatical error type and specific location; then write a script to process the transformed data, and finally get a statement without syntax errors.

4.1.2. Evaluation Index. In this study, $F_{0.5}$ [9] is used as an index to evaluate the effect of the model, which is used to measure the effect of English grammar error correction. Its calculation method is as follows: Equation (9).

$$F_{0.5} = \frac{(1 + 0.5^2) * R * P}{R + 0.5^2 * P}, \quad (16)$$

where R represents the correction rate and P represents the correction rate; they can be calculated by the following equations, respectively.

$$R = \frac{\sum_{i=1}^n |g_i \cap e_i|}{\sum_{i=1}^n |g_i|}, \quad (17)$$

$$P = \frac{\sum_{i=1}^n |g_i \cap e_i|}{\sum_{i=1}^n |e_i|}. \quad (18)$$

4.1.3. Parameter Setting. In this model test experiment, the experimental group chooses global attention mechanism +Seq2Seq as the training model, chooses word embedding for word vectorization, and adds a normalization layer. In the control group, LSTM was used to realize Seq2Seq model training, in which the hidden layer of LSTM was set to 512, the Batch size was set to 64, and the word vector dimension was 256.

4.1.4. Experimental Results and Analysis. In order to verify the error correction effect of English grammar based on Seq2Seq model proposed in this study, the proposed model is used to train the experimental data set and compared with the training results of hierarchical language model and CAMB model. The results are shown in Figure 3. As can be seen from Figure 3 is language model and CAMB model training results. Although the recall rate is slightly lower than that of the comparison model, the overall $F_{0.5}$ value is higher than that of the comparison model. Therefore, the English grammar error correction model proposed in this study has good error correction effect and obvious advantages.

At the same time, in order to verify the above English grammar error correction ability, some English sentences are selected for example testing. First, the English sentences to be corrected are divided into experimental group and control group. Among them, the experimental group is a sentence containing English grammatical errors, which is used to check whether the system can correct the sentence correctly; the control group is English grammati-

cally correct sentences, which are used to check whether the system can correct errors.

Figure 4 is the grammar correction result of the experimental group system, and Figure 5 is the error correction result of the control group system. As can be seen from Figure 4, there are many English grammatical errors in case 1 and case 2. After systematic error correction, the grammatical errors in case 1 have been corrected, but the tense error of "realized" in case 2 has not been corrected. Therefore, the error correction effect of this system model is good, but there is still room for improvement. As can be seen from Figure 5, the system does not correct the correct English sentences, which shows that the system model has good recognition ability for English grammatically correct sentences. Generally speaking, the error correction model of this system has high error correction accuracy and can be used for actual English grammar error correction.

4.2. Experiment and Result Analysis of Feedback Filtering Algorithm. In order to verify the effectiveness of the feedback filtering algorithm in this study, this algorithm is used to train grammatically correct texts in NUCLE-release2.2 corpus. First, configure srilm in the Linux system, and execute the command `n-gram-count-text train_data.txt-order3-lm model.lm` to get the trained model file `model.lm`. Then, the trained model is used to calculate the sentence probability value and confusion degree to realize feedback filtering. Usually, the probability value of syntactically incorrect statements is lower, while the probability value of syntactically correct statements is higher. Therefore, the probability value of the corrected grammatical error statement will be improved to a certain extent. Take the following statement as an example.

First, security systems have improved in many areas such as on the school campus or at the workplace.

User modifications: first, security systems have improved in many areas such as school campus or at the workplace.

Calculation result is in Figure 6.

From the above calculation results, it can be seen that the sentence confusion degree of the user and the system modification result is 171.859 and 134.366, respectively, that is, if the confusion degree of the modification result is lower, the suggestion will not be adopted. In addition, according to the grammar knowledge, the system modification method is better, so the modification proposal of the model is to perform filtering, which achieves the expected purpose.

4.3. System Application Verification

4.3.1. Framework of English Grammar Error Correction System. In order to verify the feasibility of the proposed grammar error correction again, a simple and practical English error correction website is constructed to assist English learning. The specific framework is shown in Figure 7. In this system, users can log in through account number and password and use the grammar error correction function of the system. After the user logs in to the system, input the English text to be corrected, and the system

TABLE 1: NUCLE-release2.2 dataset annotation example.

NID	PID	SID	TOKENID	TOKEN	POS	DPHEAD	DPREL	SYNT
829	1	2	0	This	DT	1	Det	(ROOT(NP*
829	1	2	1	Will	NN	7	Nsubj	*)
829	1	2	2	—	—	—	—	*
829	1	2	3	If	IN	4	Mark	(SBAR*
829	1	2	4	Not	RB	7	Dep	(FRAG*
829	1	2	5	Already	RB	4	Dep	(ADVP**))
829	1	2	6	—	—	—	—	*
829	1	9	7	Caused	VBD	-1	Root	(VP*
829	1	2	8	Problems'	NNS	7	Dobj	(NP*)

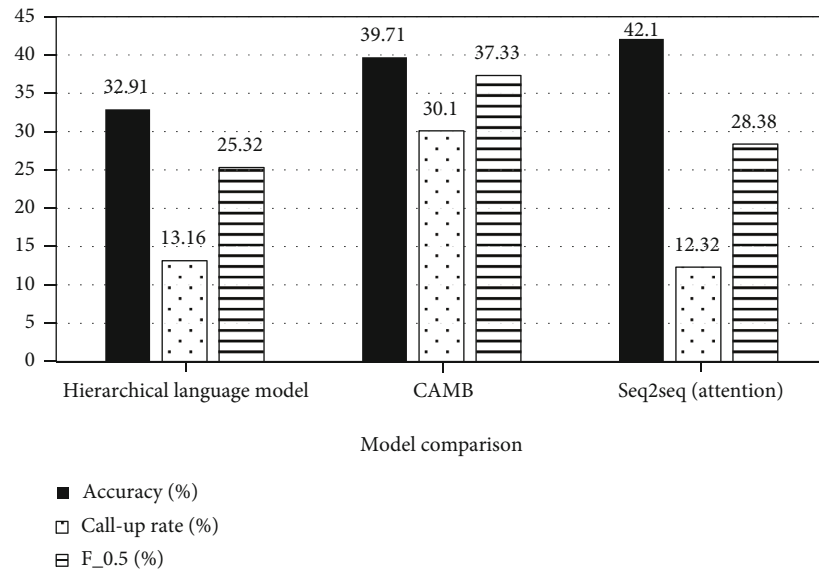


FIGURE 3: Error correction results of different models.

Case 1

Original sentence: Therefore, the equipment of biometric identification tend to be in-expensive.

System error correction: Therefore, the equipment of biometric identification tends to be inexpensive.

Case 2

Original sentence: in addition, we realized that burning of fuels produce the large amount of greenhouse gases.

System error correction: In addition, we realized that the burning of fuels produces a large amount of greenhouse gases.

FIGURE 4: Error correction results of experimental group.

corrects grammar errors according to the trained grammar error correction model, and returns the results to the user.

4.3.2. Error Correction Algorithm Training. According to the system flow, the training flow is designed as shown in Figure 8. When the system receives the English grammar error correction request, it first obtains the training data

from the original corpus and preprocesses the corpus after the corpus reaches the set threshold. Then, initialize the model and start training. Finally, the error correction effect of the model is evaluated according to the training results, and if the error correction effect of the model is improved, the grammar error correction model is updated; on the contrary, the algorithm is directly terminated. In the whole

Case 1
 Original sentence : our current population is 6 billion people and it is still growing exponentially.
 System error correction: our current populationis 6 billion people and it is still growing exponentially.

Case 2
 Original sentence : such as, global surface temperature has risen by 0.74 degree celsius in past century.
 System error correction: such as, global surface temperature has risen by 0.74 degree celsius in the past century.

FIGURE 5: Error correction results of control group.

Firstly, security systems have improved in many areas such as on the school campus or at the workplace .
 1 sentences, 20 words, 0 OOVs
 0 zeroprobs, logprob = -44.694 ppl = 134.366 pppl = 171.673

Firstly, security systems have improved in many areas such as school campus or at the workplace .
 1 sentences, 18 words, 0 OOVs
 0 zeroprobs, logprob = -42.4683 ppl = 171.859 pppl = 228.743

FIGURE 6: Filter calculation results.

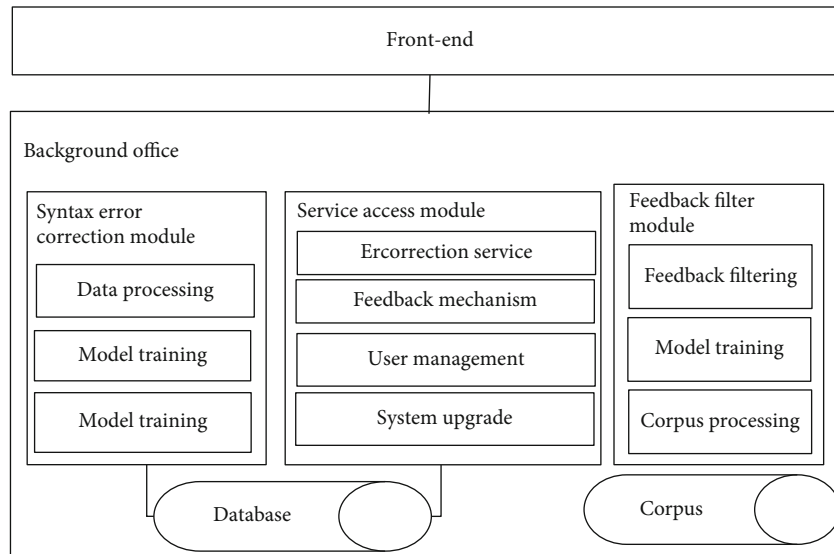


FIGURE 7: Architecture of computer-based English-assisted learning system.

process, if any link is abnormal, the system will send a training exception notification to the administrator, waiting for the system to repair.

“Original corpus” refers to the database that has been trained and can only be used as an input variable to provide rule reference for other models.

4.3.3. Implementation Process of English Grammar Error Correction. English grammar error correction module is the core functional module of this system. When the module interface receives a grammar error correction request, it must first verify whether the parameters are legal, and if

the parameters are legal, it will continue to perform grammar error correction. Then judging whether there are multiple statements in the input statement, and if it is only one statement, directly using the error correction model to correct errors; if there are multiple sentences, the error correction model should be used to correct the errors after the sentences are broken. Finally, the syntax correction can be ended by outputting the error correction result. The whole process is shown in Figure 9.

“Is the parameter legal” represents whether the entered parameter is valid, if so proceed to the next step, otherwise, end the program. “Multiple Statements” determines whether

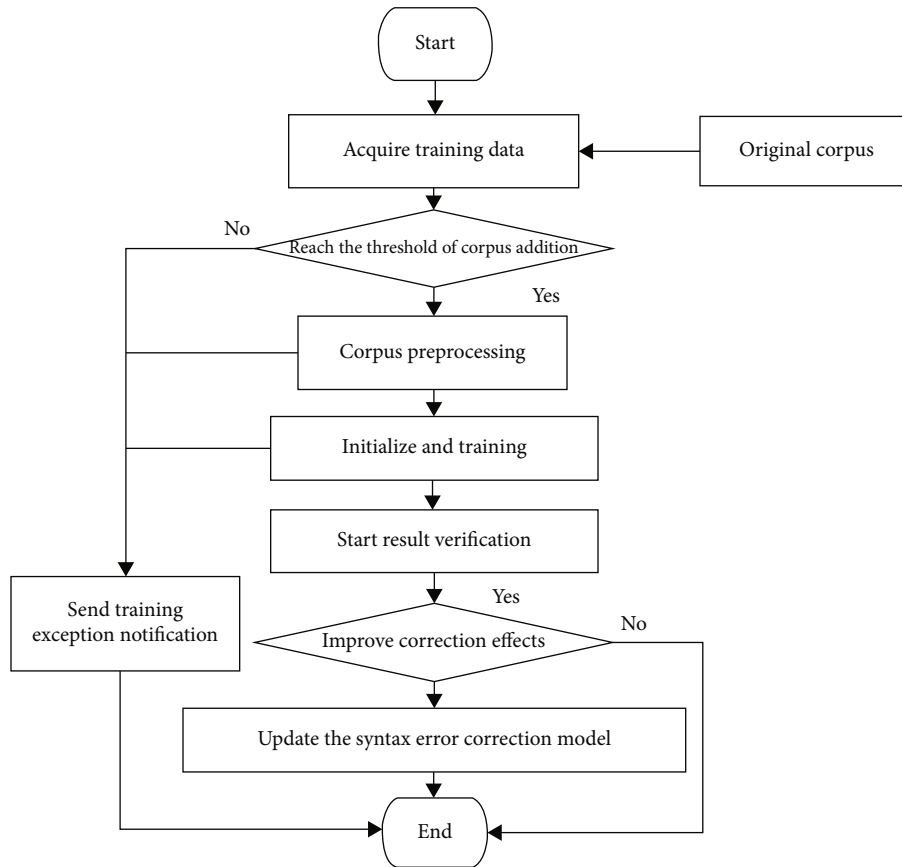


FIGURE 8: Model training flow.

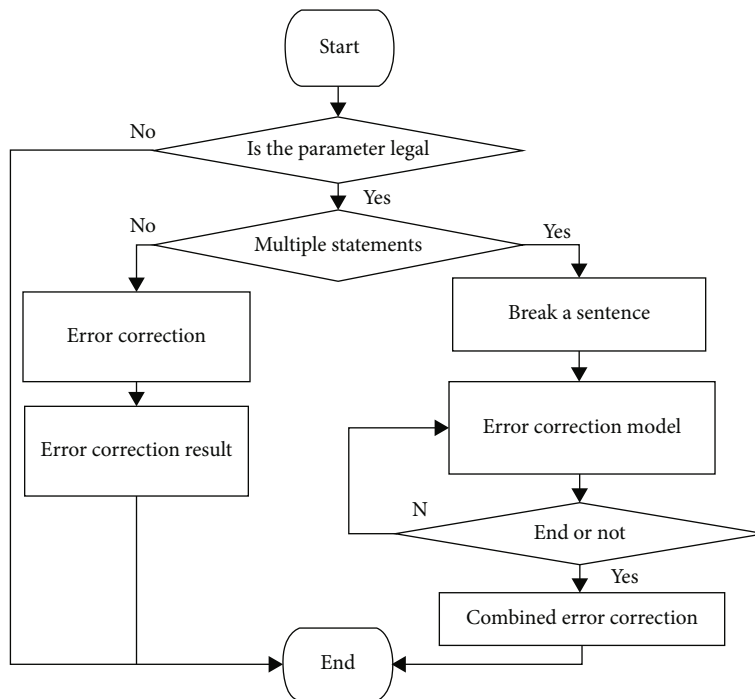


FIGURE 9: Syntax error correction flow.

FIGURE 10: Error correction page.

the complex type of statement moves on to the next implementation step.

4.3.4. System Development Environment. The front-end and back-end of this research system are developed based on react and Python languages, respectively. The related technologies used in the front end mainly include WebPack, React-Router, and Babel, while the related technologies used in the back end mainly include Django. MySQL is selected as the core database for storing data in this system, and Redis is selected as the cache database in some occasions of this system.

4.3.5. Error Correction Effect. The user logs in to the system to enter the grammar error correction interface, as shown in Figure 10. According to the prompt, enter the statement and submit it. Click the corresponding button to view the error correction result, modification suggestions, etc.

5. Conclusion

To sum up, the English grammar error correction based on **Seq2Seq** constructed in this study realizes the effective extraction of text information and improves the effect of English grammar error correction. Compared with the hierarchical language model, the training results of CAMB model show that this model has better error correction effect and higher error correction accuracy. At the same time, through feedback filtering, the system can modify the user's statement, thus improving the error correction ability. Therefore, this training model can be used to correct English grammar errors, and then assist users to learn English. However, the research work of this paper still has the problem that the error correction accuracy of complex sentences is reduced, and the complexity of the model is high. Further application of the model needs to be further improved, and more models need to be combined. The next research work needs to analyze the complexity of the model and adopt a better method to achieve it. More models are needed to verify the error correction ability of the method in this paper, and a better method is used to achieve efficient error correction efficiency.

Data Availability

The experimental data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declared that they have no conflicts of interest regarding this work.

References

- [1] G. Kaur, K. Kaur, and P. Singh, "Spell checker for Punjabi language using deep neural network," in *2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS)*, pp. 147–151, Coimbatore, India, 2019.
- [2] L. U. Yuqing, Y. Hong, L. U. Jun, J. Yao, and Q. Zhu, "Context dependent word error detection and correction," *Journal of Chinese Information Processing*, vol. 25, no. 1, pp. 85–90, 2011.
- [3] C. Wang, L. Y. Xu, and J. S. Fan, "A general deep learning framework for history-dependent response prediction based on UA-Seq2Seq model," *Computer Methods in Applied Mechanics and Engineering*, vol. 372, article 113357, 2020.
- [4] J. Lee and D. W. Kim, "Memetic feature selection algorithm for multi-label classification," *Information Sciences*, vol. 293, pp. 80–96, 2015.
- [5] D. Kwon, K. Natarajan, S. C. Suh, H. Kim, and J. Kim, "An empirical study on network anomaly detection using convolutional neural networks," in *IEEE International Conference on Distributed Computing Systems*, pp. 1595–1598, Vienna, Austria, 2018.
- [6] D. Jeong, M. Baek, and S. S. Lee, "Long-term prediction of vehicle trajectory based on a deep neural network," in *International Conference on Information & Communication Technology Convergence*, pp. 725–727, Jeju, Korea (South), 2017.
- [7] L. You, S. Xiao, Q. Peng, C. Claramunt, and J. Zhang, "ST-Seq2-Seq: a spatio-temporal feature-optimized Seq2Seq model for short-term vessel trajectory prediction," *IEEE Access*, vol. 8, pp. 218565–218574, 2020.
- [8] Y. Kuriyagawa, H.-E. Im, I. Kageyama, and S. Onishi, "A research on analytical method of driver-vehicle-environment system for construction of intelligent driver support system," *Vehicle System Dynamics*, vol. 37, no. 5, pp. 339–358, 2002.
- [9] G. Samson, "From writing to doing: the challenges of implementing integration (and interdisciplinarity) in the teaching of mathematics, sciences, and technology," *Canadian Journal of Science Mathematics and Technology Education*, vol. 14, no. 4, pp. 346–358, 2014.