

Research Article

A Novel Joint Entity Relation Extraction Based on Capsule Network and Part-of-Speech Weighting

Jianmin Wang,¹ Yujia Song,¹ Wenbin Zhao ,¹ Ziyue Jia,¹ and Feng Wu²

¹School of Information Science and Technology, Shijiazhuang Tiedao University, Shijiazhuang, Hebei, China

²Hebei Science and Technology Information Processing Laboratory, Hebei Institute of Science and Technology Information, Shijiazhuang, Hebei, China

Correspondence should be addressed to Wenbin Zhao; zhaowb2013@stdu.edu.cn

Received 14 May 2022; Revised 28 June 2022; Accepted 7 July 2022; Published 9 August 2022

Academic Editor: Imran Shafique Ansari

Copyright © 2022 Jianmin Wang et al. 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.

With the development of science and technology, science and technology policies are increasing year by year. Science and technology policies are literature existing in the form of texts, which are characterized by rigorous structure, clear hierarchy, and standard language. Mining template information from policies can optimize data templates and improve the efficiency of recommending data to users. This paper proposes a joint entity relation extraction model based on capsule networks and part-of-speech weighting. In order to learn more feature information from word vector, capsule network based on bidirectional gated cyclic unit is used to replace the traditional convolutional neural network. In view of the phenomenon of imperfect semantic expression of word vector, part-of-speech features are added to enrich text information. Meanwhile, in order to solve the weight distribution problem of word features and part-of-speech features, an artificial fish swarm algorithm is proposed to optimize the two feature weights by iterative optimization, and the effectiveness of the proposed model is proved by experiments.

1. Introduction

Big data is gradually becoming a new model and element to promote the construction of social development and a modern governance system, which has had a huge impact on national governance, science and technology decision-making, and intelligence services. It has become one of the consensus [1] in the field of intelligence services research to effectively make use of massive data resources and big data analysis methods and technologies to support decision makers. Science and technology policy is a document in the form of text, which has the characteristics of rigorous structure, clear hierarchy, and standard language. Mining template information from the policy can optimize the data template and improve the efficiency of recommending data to users. Through the extract the entity relation from science and technology policy, the theme of data stored in the form of template database not only save time and manpower and can promote the efficient utilization of data and optimize the user of data viewing experience, so that the support project

research direction and the formulation of science and technology policy is of great significance.

Entity relation extraction refers to the extraction of relation triples (entity 1, entity 2, and relation) from sentences according to context semantics, which is one of the important subtasks of relation extraction [2]. It is widely used in text summarization, knowledge graph, and search engine. Relation extraction can be divided into two types according to the set of relation extraction results: restricted relation extraction, which requires the set of relation extraction to be determined in advance, and which category the current relation between entities belongs to according to entity and context semantics, similar to text classification; open relation extraction, there is no predetermined relation set and the expected domain of extraction is uncertain.

In this paper, a joint entity relation extraction model of capsule network based on bidirectional gated cyclic unit generation is constructed. Experiments show that the macro-average F1 value based on BLSTM model is higher than that based on BGRU model, and the running time

based on BGRU model is shorter than that based on BLSTM model. Considering the imperfect semantic expression of word vector and the location information contained in the capsule network, two part-of-speech weighted models based on self_attention were constructed by adding part-of-speech features to enrich the text information, and the validity of the later combined model was verified by experiments. Aiming at the problem of weight distribution of word feature and part-of-speech feature, artificial fish swarm algorithm is proposed to optimize the weight of the two features by iterative optimization to improve the classification effect. The effectiveness of the two optimization schemes is proved by experiments.

2. Related Work

2.1. Entity Relation Extraction. Deep learning supervised entity-based relation mainly includes pipe lined extraction and joint extraction [2]. Pipe lining extraction refers to the realization of named entity recognition, NER, and then judging the relation between the entities, the benefit of this way is very convenient to handle and flexible in combination, but it ignores the association between the two subtasks and produces false superposition [3]. Joint extraction builds entity recognition and relation extraction within a model, which improves the extraction effect by enhancing the dependence of the two subtasks by sharing the encoding layer. Miwa and Bansal [4] used neural networks for joint relation extraction to improve the accuracy of the model, which is the first joint extraction model of neural networks. The joint extraction model designed by Li et al. [5] contains three layers: word embedding, sequence, and dependency layers. The proposed model obtain good results on the semeval-2008 dataset. Zheng et al. [6] uses Bi-LSTM to extract text semantic information at the encoding layer as a shared layer for entity recognition and relation extraction, and enter data from the coding layer into the entity recognition model and relation extraction model. The authors experiment the model on the public dataset ACE05 to the highest level and the effectiveness of the joint extraction model is proved. Zhang et al. [7] combine convolutional neural networks with support vector machines and conditional random airports. A joint neural network model is constructed and parameter sharing way in the drug manual corpus has achieved good results. Ma et al. [8] put forward the entity and relation joint extraction model of the feedback mechanism. A feedback mechanism is introduced to modify the feature extraction layer and entity recognition model parameters according to the weighted loss and feedback of the two loss values.

2.2. Capsule Network. Capsule network is a variant of Convolutional Neural Networks, which was first proposed by Geoffrey Hinton et al. The capsule network uses the capsule instead of the neuron node of the traditional neural network. The capsule is a neuron vector and each value of the capsule is a vector, which contains the feature, direction, state, and other information of the entity. The relation between words is trained by dynamic routing. Then the

feature vectors are clustered to capture the text space features, including input layer, convolution layer, main capsule layer, category capsule layer, and output layer. Sabour et al. [9] built the capsule network architecture surpassed CNN on the MNIST data set to achieve the most advanced performance at that time. Wang et al. [10] and others designed a hybrid depth network model based on capsule network and cyclic neural network to verify the effectiveness of the model compared with the popular text classification methods at present. The author shows that the location information of the text and the relation between the part and the whole can be learned through the capsule network, which can enrich semantics and reduce the loss of feature information. The integrated model based on transformer-capsule proposed by Tang et al. [11] proves that the capsule network can effectively extract the phrase features in the text.

3. Joint Entity Relation Extraction Model Based on Capsule Network and Part-of-Speech Weighting

3.1. Entity Relation Extraction Based on Capsule Network. This section constructs a parameter-shared joint entity relation extraction model based on capsule network. As shown in Figure 1, the model is mainly composed of four parts: vector representation layer, feature learning layer, entity recognition layer, and relation classification layer.

- (1) Vector representation layer. For the problem that the computer cannot deal with the text directly, the word is embedded into the vector to represent the text feature, and word2vec [12] is used to train the domain word vector to enhance the semantic accuracy.
- (2) Feature learning layer. Word vectors can only represent the general meaning of the text. BGRU [13] is used to learn higher level contextual semantic features from the word vector representation of sentences.
- (3) Weight adjustment layer. For the learning text, feature weights are unreasonable and cannot make efficient use of resources, and the self-attention mechanism [14] is used to calculate weights as the weight adjustment layer. Vector representation layer, feature learning layer, and weight adjustment layer are used as the shared part of the model.
- (4) Entity recognition layer. Conditional random field is used to label text sequentially. Conditional random field is a typical discriminant undirected probability graph model, which was proposed by Lafferty [15] in 2001, which combines the characteristics of hidden Markov model (Hidden Markov Model, HMM) [16] and maximum entropy model (Maximum Entropy Markov Model, MEMM) [17], but the conditional random field does not have two unreasonable hypotheses of hidden Markov model (homogeneous Markov hypothesis and observation independence hypothesis). This makes the algorithm more flexible

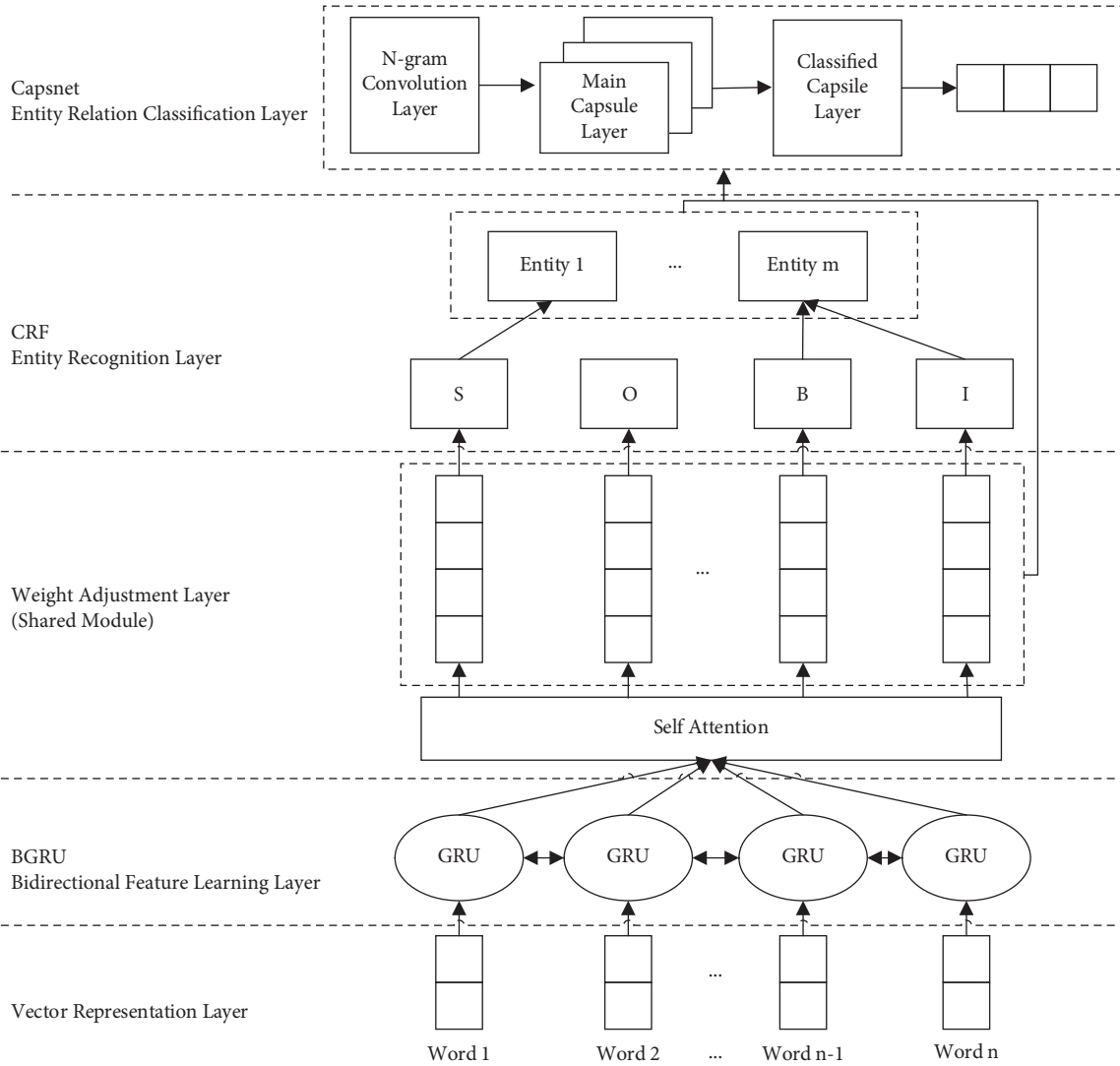


FIGURE 1: Joint entity relation extraction model based on capsule network.

in design with more contextual information. Compared with the maximum entropy model, the conditional random field calculates the global optimal node probability, not just the local normalization, so the problem of label bias is solved. McCallum [18] took the lead in using it in named entity recognition, and after continuous improvement, it has become the most successful method in named entity recognition [19]. According to the BIOS tagging method, the text is marked with a corresponding tag, in which the words connected by *B* and *I* are combined as an entity, and there is only one *B* tag as the beginning of the entity, and a plurality of *I* tags can be connected continuously in the middle; the *O* tag identifies that the recognition result of the current word does not belong to any entity and does not be processed; the *S* tag can be used as a separate entity, and when there are continuous *S* tags, it is recognized as a plurality of entities.

(5) Relational classification layer. The capsule network uses the capsule instead of the neuron node of the traditional neural network. The capsule is a neuron vector and each value of the capsule is a vector, which contains the feature, direction, state and other information of the entity. The relation between words is trained by dynamic routing, and then the feature vectors are clustered to capture the text space features.

The first layer is the input layer, in which the weight-adjusted feature coding layer and the recognized entity vector are spliced as the input of the capsule network $ABSR = [absr_1, \dots, absr_i, \dots, absr_L]$.

The second layer is the convolution layer, which extracts sentence features through standard *N*-gram convolution, in which is the size of the convolution kernel determines the quality of feature extraction. The input *ABSR* of the capsule network layer is $L * V$ dimensional data, where *L* is the data length after the stitching of

sentence features and entity vectors and V is the embedded word vector dimension. Set the filter with a window size of K and convolution sentences with a step size of 1, as shown in the following formula:

$$m_i^a = f(z_{i:i+(K-1)} \circ M^a + b_0). \quad (1)$$

Among them, $M^a \in M^{K \times V}$ represents filter for convolution operation, f represents the activation function ReLU, b_0 represents an offset item, $z_{i:i+(K-1)}$ represents the data in the range of time i window from sentence i to $i+K-1$, you will eventually get a feature sequence with $L-K+1$ length $\text{ABSR} = [\text{absr}_1, \dots, \text{absr}_i, \dots, \text{absr}_{L-(K+1)}]$.

The third floor is PrimaryCaps, the data extracted by the convolution layer is still a scalar representation. The main capsule layer converts the features extracted by the N -gram convolution layer into capsules and retains the instantiated parameters of the features in the form of vectors. In the shared parameter mode, each N -gram vector is multiplied by the shared parameters, and then transformed into vector neurons for dynamic routing.

The fourth floor is ClassCaps, the dynamic routing algorithm is used to complete the transformation from the main capsule layer to the class capsule layer. The dynamic band routing algorithm ensures that the information of the child capsule is sent to the corresponding parent capsule in a nonlinear mapping way. Through several iterations to enhance or weaken the connection strength between the child capsule and the parent capsule, dynamic routing is more effective than the maximum pooling of CNN to lose a lot of text space information [20]. In the category capsule layer, the module operation is done for each vector, and the vector with the maximum modulus indicates that the more likely the analogy is.

Dynamic routing is the main part of the algorithm and the computing model is shown in Figure 2.

In Figure 2, b_{ij} represents the j weight of the i th cycle, the initial default is 0, v_1 and v_2 are the lower eigenvectors output from the upper capsule layer, W is the transformation matrix corresponding to the vector, r is the number of iterations, and c_i indicates the similarity between the upper capsule and the lower capsule. By updating each iteration, squeeze the vector using the Squash function as the activation function, and the calculation formula is as shown in formula (2). The larger the module of the vector is, the stronger the feature is. After the end of the calculation, the b value is updated to start the next round of calculation until the iteration r stops.

$$\text{Squash}(x) = \frac{\|x\|^2}{1 + \|x\|^2} \frac{x}{\|x\|}. \quad (2)$$

The fifth floor is the output layer. The output layer uses flattening function plus full-link layer plus softmax function. The output of the capsule network is cap_num^* and cap_dim dimensional data, where cap_num represents the number of capsules and cap_dim represents the capsule dimension. You need to flatten the output value of the capsule network into one-dimensional data and transfer it to the full-connection layer.

The full-connection layer plays the role of classifier in the neural network, which belongs to the feedforward neural network. The input of the full-connection layer is all the output from the upper layer, and the output is n output nodes, where n represents the number of categories. The output of the full-connection layer is input into the softmax function to calculate the probability, and the final classification result is obtained.

3.2. Entity Relation Extraction Based on Part-of-Speech Weighting. Part of speech can not only improve the accuracy of entity recognition but also promote the problem of text classification. Considering that the capsule network contains location information, this chapter proposes to add part of speech to reduce the ambiguity of words and enrich the feature semantic information. In addition, considering that part-of-speech features play different roles in different data, part of speech and other features are mostly added to the model of text data processing to deal with part of speech and other features in the same proportion. However, from the point of view of the different effects of part of speech on different experiments, the importance of part of speech plays a different role in different data and different models, so a joint entity relation extraction model based on artificial fish swarm part-of-speech weighting is constructed. The artificial fish swarm algorithm is used to optimize the weight of word features and part-of-speech features by iterative optimization, so as to improve the classification effect [21].

The division of part of speech is based on grammatical norms and the meaning of words. Learning the context features of part of speech through BGRU can increase the grammatical information of sentence vectors. Considering that parts of speech have different importance to features, self-attention is used to add weight to each part of speech. As shown in Figure 3, the model consists of six parts: embedding layer, weight adjustment layer 1, feature learning layer, weight adjustment layer 2, feature binding layer, and relation extraction layer.

Embedded layer. According to the word vector dictionary and the part-of-speech vector dictionary, the word and part of speech are transformed into vector data by word embedding technology, $S = [w_1, \dots, w_i, \dots, w_n]$ and $P = [p_1, \dots, p_i, \dots, p_n]$, the weight of the artificial fish school $Q = q$ directly as input data.

Weight adjustment layer 1. In the weight adjustment layer 1, the part-of-speech vector is weighted according to the weight of the artificial fish school, and the weight of the artificial fish school is multiplied by the part-of-speech vector, so as to adjust the weight of the part-of-speech vector and the part-of-speech vector, $QP = [qp_1, \dots, qp_i, \dots, qp_n]$.

Feature learning layer. Feature learning layer requires a context semantic and meaning vector of the word vector, a context whispering, $BS = [bw_1, \dots, bw_i, \dots, bw_n]$ and $BQP = [aqp_1, \dots, bqpi, \dots, bqpn]$, at this time, BQP vector is indicated by the weighted word semantic.

Weight adjustment layer 2. Weight adjustment layer 2 needs to adjust the internal weight of each word and part of speech in the word vector and part-of-speech vector,

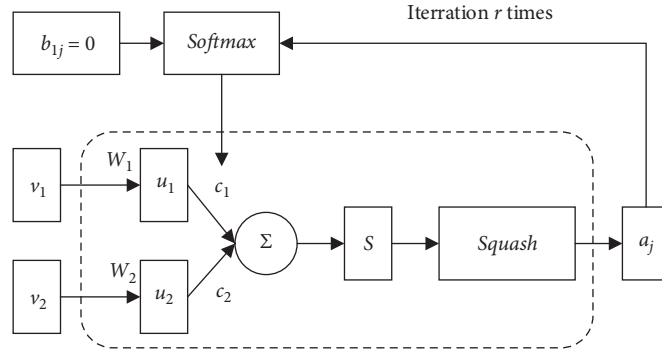


FIGURE 2: Dynamic routing model of capsule network.

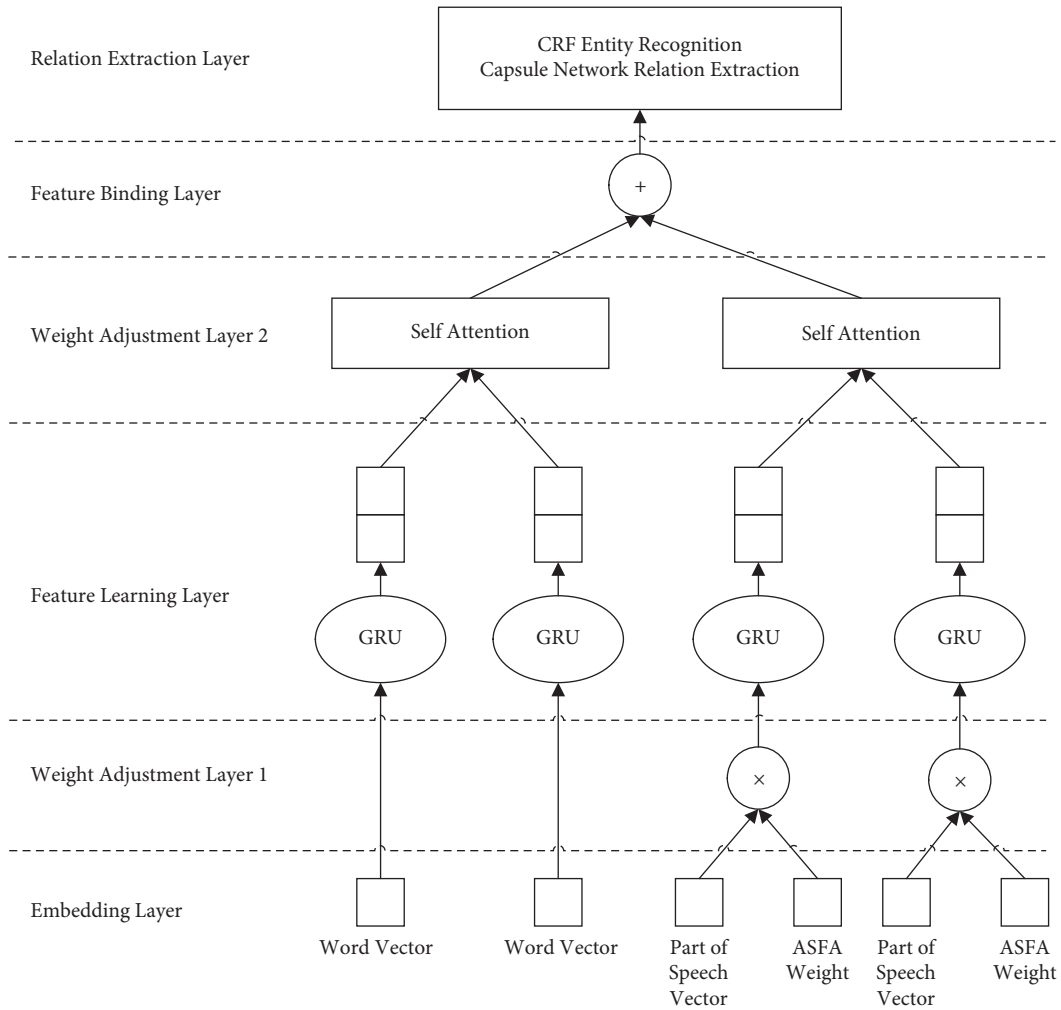


FIGURE 3: Joint entity relation extraction model of capsule network combined with part-of-speech weighting.

respectively, through self_attention to get the word vector and part-of-speech vector representation $ABS = [abw_1, \dots, abw_i, \dots, abw_n]$, $ABQP = [abqp_1, \dots, abqp_i, \dots, abqp_n]$, the word vector, and the morphinal vector at this time are two weighted results, respectively.

Weight binding layer. The feature binding layer combines the weighted word vector with context semantics and the part-of-speech vector, $ABS = [abw_1, \dots, abw_i, \dots, abw_n]$ and

$ABQP = [abqp_1, \dots, abqp_i, \dots, abqp_n]$. After splicing, we get the word vector $ABSQP = [abwqp_1, \dots, abwqp_i, \dots, abwqp_n]$ as the output of the feature binding layer, the output vector is the semantic weighting of the context of the containing word feature and the semantic weighting of the part-of-speech feature context.

Different from the general functional model with fixed input and output, the output result of the deep learning

model is random, and different results can be obtained even if the same data is used, so the parameters of each layer in the experiment are fixed by saving the model, so that the experimental results of the model under the same data are the same, and that the deep learning can be applied to the artificial fish swarm algorithm. The flowchart of part-of-speech weight optimization based on artificial fish school is as follows:

Step 1: Artificial fish swarm algorithm needs to adjust the position of artificial fish according to the objective function, so as to get the optimal solution. However, the randomness of the parameters in the deep learning model leads to the randomness of the experimental results, so the experimental results are not accurate, and the experiment itself will fluctuate under the same configuration parameters and the same experimental data, which may cause artificial fish to fluctuate back and forth, and interfere with the optimization of artificial fish. In order to reduce the influence of randomness on the experiment, the part-of-speech weight is set to 1 into the capsule network entity relation extraction model combined with part-of-speech weighting. After 50 iterations, the trained model is saved to the parameters in the fixed depth model.

Step 2: Initialize the population size, visual field, step size, crowding factor, repetition times, and the initial position of each artificial fish.

Step 3: Through the model saved in step 1, the macro-average F1 value of each artificial fish is calculated as the objective function value of the artificial fish, and the optimal value is given to the bulletin board.

Step 4: According to the objective function value of each individual, the individual artificial fish are carried out foraging, clustering, rear-end collision, and other behaviors, by comparing with the bulletin board to choose whether to update the bulletin board.

Step 5: Judge whether the condition for the end is met, then jump out of the cycle, and the value of the bulletin board is the final result; if not, you will turn to step 3.

The better value obtained by the artificial fish swarm algorithm is used as the weight of part of speech, and the weighted part-of-speech vector is obtained by multiplying the weight value with the part-of-speech vector. The weighted part-of-speech vector and the part-of-speech vector spliced after the weighted attention mechanism are taken as the joint sharing part of the capsule network entity relation extraction model with part-of-speech weighting.

4. Experiments

4.1. Entity Relation Extraction Based on Capsule Network. In order to verify the effectiveness of the capsule network based on self-attention mechanism in joint entity relation extraction model, take the macro-average F1 value as the measurement standard. Because the initial values of many parameters are randomly generated during deep learning, most of the experiments are random, so the experimental

results are different even in the same model with the same experimental data, in order to reduce the interference of experimental randomness to the experiment. Each experiment is run three times and the average value of three experiments is taken as the final macro-average F1 experimental results. The model proposed in this chapter is compared from two aspects of model comparison and affecting model parameters by using single variable method to verify the effectiveness of the proposed method.

4.1.1. Model Structure. In order to verify the effectiveness of the proposed model, the joint entity relation extraction model based on capsule network (represented by Capsnet), the joint entity relation extraction model based on CNN (represented by CNN), and the joint extraction model of capsule network with self-attention mechanism (represented by Capsnet + self_attention) were tested with the same data. First of all, Capsnet and CNN are iterated 300 times, and it is found that the macro-average F1 value of the two model experiments fluctuates in a small range after 100 iterations, and the experimental results are shown in Figure 4, so the number of iterations is set to 100 in the future experiment.

As can be seen from Figure 4, the CNN model is small in the case where the number of iterations is small, but there is a crossover at the 9th iteration. The average value of the experimental results of the two models is from 100 to 300 iterations, respectively, the average of the CAPSNET model is approximately 0.78293 (retaining five effective numbers), the average of the CNN model has an average F1 value of 0.77489. Keeping the five effective numbers, the CAPSNET model is about 0.00804 in the CNN model.

As can be seen from Figure 5, the difference between the CAPSNET model and the CAPSNET model after the addition of the self-focus mechanism is not very large, but the model from the 20th iteration begins to add self-focus, which is slightly higher or about equal to the CAPSNET model. Occasionally, the average value of the average F1 value is from 40th to 100 iterations, using the CAPSNET as 0.77452, the model of the self-focus mechanism is 0.78043. It is 0.00591, which is 0.0051, proves that the self-focus mechanism improves the average F1 value of the model to a certain extent, which proves the effectiveness of the incoming and introduction of the capsule network entity relation of the introduction of self-focus mechanism.

4.1.2. Impact Model Parameters. In the depth learning model, suitable parameters can not only improve the operating rate of the model but also improve the learning efficiency of the model. The effects of capsule dimensions and convolutionary window values in the capsule network were studied using a single variable.

4.2. Capsule Network Based on Self-Attention Part-of-Speech Weighting. The model is analysed from two aspects: macro-average F1 value and experimental time. The details are as follows.

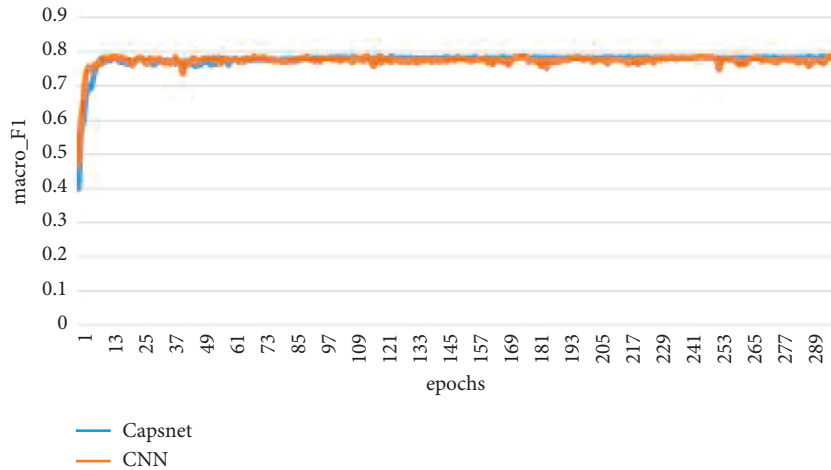


FIGURE 4: Capsnet model and CNN macro-average F1 value with the number of iterations.

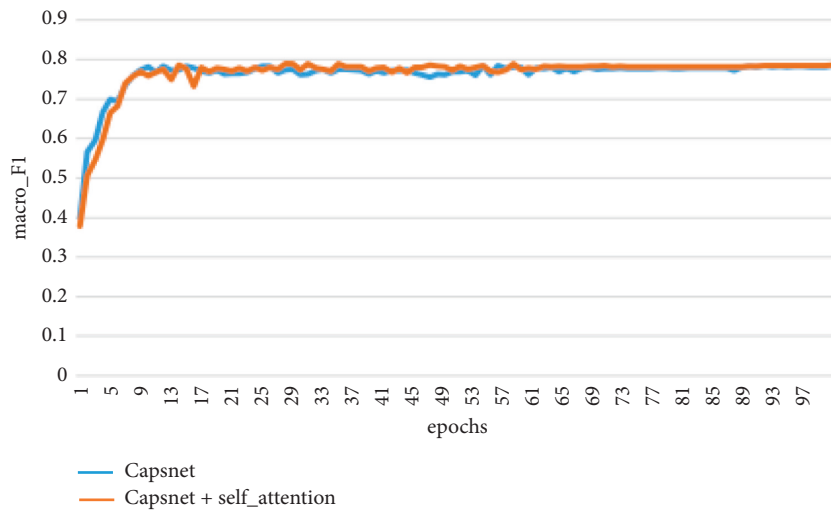


FIGURE 5: Three model macro-average F1 values with the number of iterative times.

4.2.1. Comparison of Macro-Average F1 Values of Four Models. Comparing the precombination model and post-combination model proposed by combining the features of part of speech in this section with the third chapter model (expressed by BLSTM) and the model proposed in the first Section 3 (expressed by BGRU), 50 iterations will be realized, and the experimental results are shown in Figure 8.

The macro-average F1 value of the four models from the 40th iteration to the 50th iteration was calculated, respectively, and the order of the macro-average F1 value of the four models from big to small was: late combination model 0.78394, BLSTM model 0.77792, early combination model 0.77632, and BGRU model 0.7733. It can be seen that the combination model in the early stage is 0.00302 higher than BGRU model, but it is lower than BLSTM alone. The combination model in the later stage has a great influence on the model, which is 0.01064 higher than BGRU alone and 0.00602 higher than BLSTM alone. It verifies the effectiveness of the joint entity relation extraction model in capsule network based on self_attention part-of-speech weighting.

4.2.2. Comparison of Running Time of Four Models. The four models are iterated 30 times, 50 times, and 100 times, respectively, to get the running time. The experimental results are shown in Figure 9. As far as the three experimental results are concerned, the running time of the preintegration model and the postintegration model is not much different, which is longer than that of BGRU alone, but the running time of the preintegration model and the postintegration model is slightly shorter than that of BLSTM alone.

By replacing BLSTM with BGRU and taking part of speech as an extended feature, the macro-average F1 value of the joint entity relation extraction model based on capsule network can be improved on the premise of reducing the running time, thus proving the effectiveness of the joint entity relation extraction model based on self_attention part-of-speech weighting.

4.3. Capsule Network Based on Part-of-Speech Weighting of Artificial Fish Swarm. In view of the randomness of the deep learning model, a set of data cannot represent the overall

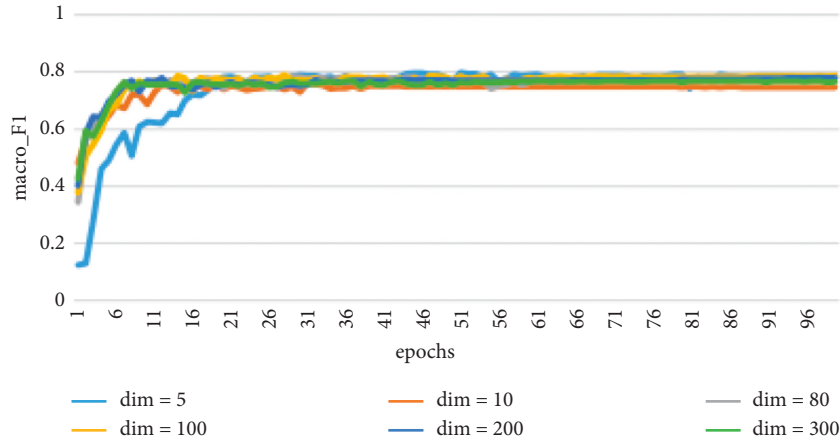


FIGURE 6: Variation of macro-average F1 value of model with iteration times under different capsule dimensions.

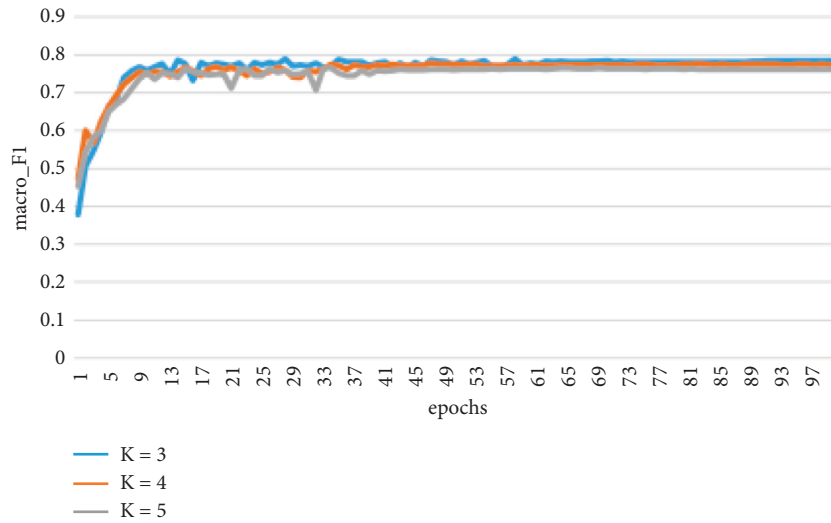


FIGURE 7: The graph of the model macro-average F1 value changing with iteration times under different convolution windows.

effect of the model. The model is iterated 50 times to save the experimental model and run three times randomly to save three models: model 1, model 2, model 3; artificial fish experiments are carried out with the same three groups of data for each model; data 1, data 2, and data 3. A total of 9 experimental results were obtained in 3 groups, each of which corresponds to the optimal value of a model and a set of data, and the average value of the 9 experimental results is 0.84293.

The experimental results of the capsule network joint entity relation extraction model without adding part of speech are shown in Figure 10. The weight of the part-of-speech feature is 0.8 (the approximate value is obtained by rounding) by the artificial fish swarm experiment, which is used to compare following models, the capsule network joint entity relation extraction model based on the artificial fish school part-of-speech weighting (represented by artificial fish school + self-attention mechanism), a POS-weighted capsule network joint entity relation extraction model based on self_attention POS

weighting (represented by self-attention mechanism), and the Pipeline model (represented by Pipeline) [22]. The final experimental results are shown in Figure 10.

Each experiment was performed three times average, and the average F1 value of the Pipeline model was 0.77043, the unrelated capsule network entity relation combined withdrawal model macro-average F1 value is 0.7733, alone uses a self-focused mechanism's postcombination model of the model. The average F1 value was 0.78394 and the average F1 value of the artificial fish group + self-focus mechanism model reached is 0.78565. It can be seen that the average F1 value of the Pipeline model ratio of the capsule network entity is increased by 0.00287; the model after joining the morphism is increased by 0.01064 than the capsule network entity relation, and the combination of artificial fish group is used in combination. The solid relational combined extraction model is 0.00171, and the macro-average F1 value of the combination of the combination of the combination of self-focusing mechanisms is 0.00171, which proves the effectiveness of the model.

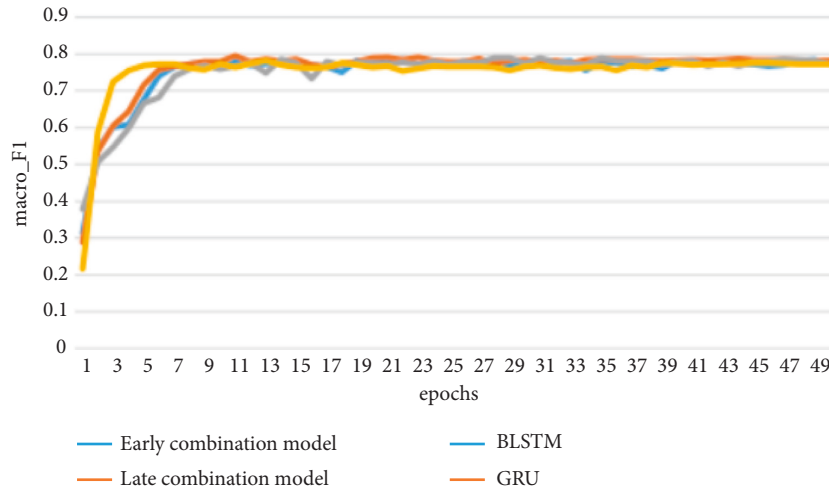


FIGURE 8: Diagram of macro-average F1 value of four models changing with iteration times.

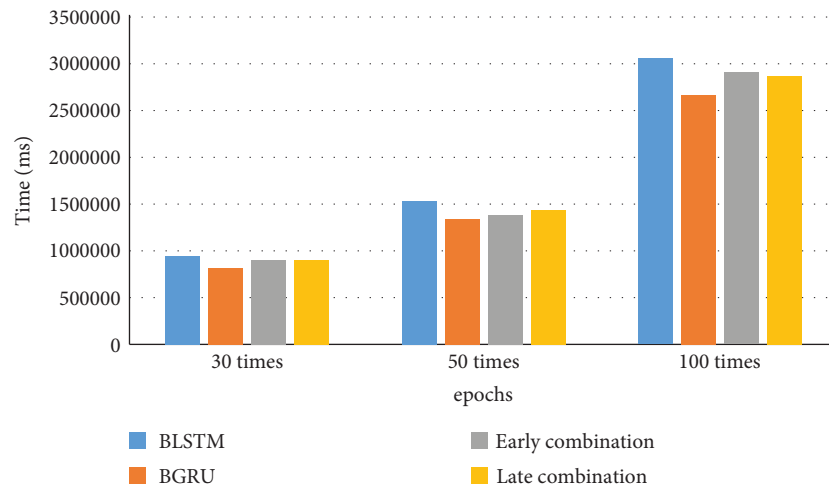


FIGURE 9: The running time of four models varies with the number of iterations.

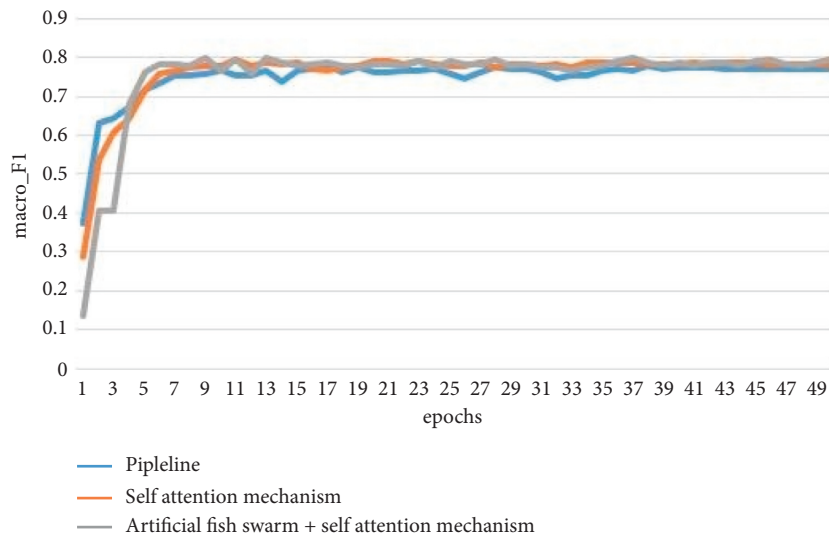


FIGURE 10: Variation of macro-average F1 with different feature inputs.

5. Conclusion

On the basis of the capsule network joint entity relation extraction model, the semantic information is improved by adding part-of-speech features and an entity relation joint extraction model based on self_attention is constructed to weigh the part of speech in the sentence internally. The experimental results show that the macro-average F1 value of the model is improved. In order to solve the problem of the weight distribution of word features and part-of-speech features, an artificial fish swarm algorithm is proposed to optimize the two feature weights by iterative optimization, and the classification effect is improved by adjusting the weight of part of speech to control the proportion of word vectors and part-of-speech vectors. To sum up, the part of speech can improve the features of part of speech to a certain extent and the importance of part of speech is different from that of words. In the future, semantic feature representation will be improved by enriching semantic feature, perfecting model structure and optimizing function method.

Data Availability

The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that there are no conflicts of interest with any financial organizations regarding the material reported in this article.

Acknowledgments

The authors acknowledge the National Natural Science Foundation of China (61373160), the research project "Research on Recognition Method of Knowledge Evolution Path for Sequential Associated Text Based on Graph Neural Network (F2021210003)" of the Natural Science Foundation of Hebei Province, the research project "Knowledge Graph Construction of Multi-Source Domain Data Based on Knowledge Representation Learning (QN2020197)" of the Education Department of Hebei Province, and the research project of Hebei Science and Technology Information Processing Laboratory.

References

- [1] Q. Zhou and H. B. Lin, "Research on Topic Recognition of Science and Technology Decision Requirements Based on Multi-Source Data Fusion," *Information Theory and practice*, vol. 42, no. 3, pp. 107–113, 2019.
- [2] E. Haihong, W. Zhang, S. Xiao et al., "A review of in-depth learning entity relation extraction," *Journal of Software*, vol. 30, no. 6, pp. 1793–1818, 2019.
- [3] T. Xia and Y. Sun, "A study on the extraction method of Tibetan solid body based on the combined model," *Chinese Journal of Information and Information*, vol. 32, no. 12, pp. 76–83, 2018.
- [4] M. Miwa and M. Bansal, "End-to-end relation extraction using LSTMs on sequences and tree structures," in *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, pp. 1105–1116, Berlin, Germany, August 2016.
- [5] F. Li, M. Zhang, G. Fu, and D. Ji, "A neural joint model for entity and relation extraction from biomedical text," *BMC Bioinformatics*, vol. 18, no. 1, 198 pages, 2017.
- [6] S. Zheng, Y. Hao, D. Lu et al., "Joint entity and relation extraction based on a hybrid neural network," *Neurocomputing*, vol. 257, pp. 59–66, 2017.
- [7] Y. Zhang, M. Liu, and H. Hu, "Chinese medical entity classification and relationship extraction based on joint neural network model," *Computer Engineering and Science*, vol. 41, no. 6, pp. 1110–1118, 2019.
- [8] J. Ma, Z. Li, H. Zhu, and W. Zikai, "Frieeces and relationships of feedback mechanisms joint extraction method," *Computer Science*, vol. 46, no. 12, pp. 242–249, 2019.
- [9] S. Sabour, N. Frosst, and G. E. Hinton, "Dynamic routing between capsules," in *Annual Conference on Neural Information Processing Systems*, pp. 3856–3866, 2017.
- [10] H. Wang, W. Liu, and K. Guo, "A capsule network text classification method based on mixed word vector," *Small microcomputer system*, vol. 41, no. 1, pp. 218–224, 2020.
- [11] Z. Tang, Z. Wang, A. Zhou, F. Meishan, Q. Wen, and M. Lu, "Transformer-capsule integration model for text classification," *Computer Engineering and Application*, vol. 56, no. 24, pp. 151–156, 2020.
- [12] T. Mikolov, I. Sutskever, C. Chen, G. S. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," *Advances in Neural Information Processing Systems*, vol. 26, pp. 3111–3119, 2013.
- [13] K. Cho, B. Van Merriënboer, Ç. Gülçehre et al., "Learning Phrase Representations Using RNN Encoder-Decoder for Statistical Machine Translation," in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2014.
- [14] D. Bahdanau, K. Cho, and Y. Bengio, "Neural Machine Translation by Jointly Learning to Align and Translate," in *International Conference on Learning Representations*, 2014.
- [15] K. L. Anh and T. S. Nguyen, "Effective Combination of Bert Model and Cross-Sentence Contexts in Aspect Extraction," in *Proceedings of the International Conference On Machine Learning*, Chongqing, China, July 2021.
- [16] S. Sadhu and H. Hermansky, "Continual learning in automatic speech recognition," in *Proceedings of the Interspeech*, pp. 1246–1250, Shanghai, China, October 2020.
- [17] M. Peng, R. Ma, Q. Zhang, and L. Zhao, "Toward Recognizing More Entity Types in NER: An Efficient Implementation Using Only Entity Lexicons," in *Findings of the Association for Computational Linguistics: EMNLP*, pp. 678–688, 2020.
- [18] Y. Nie, Y. Tian, Y. Song, and A. Xiang, "Improving Named Entity Recognition with Attentive Ensemble of Syntactic Information," in *Findings of the Association for Computational Linguistics: EMNLP*, pp. 4231–4245, 2020.
- [19] R. Dror, L. Peled-Cohen, S. Shlomov, and R. Reichart, "Statistical significance testing for natural language processing," *Synthesis Lectures on Human Language Technologies*, vol. 13, no. 2, pp. 1–116, 2020.
- [20] W. Zhao, J. Ye, and M. Yang, "Investigating capsule networks with dynamic routing for text classification," in *Proceedings of*

the 2018 Conference on Empirical Methods in Natural Language Processing, Brussels, Belgium, October 2018.

- [21] N. M. Sureja and S. P. Patel, "Solving A combinatorial optimization problem using artificial fish swarm algorithm," *International Journal of Engineering Trends and Technology*, vol. 68, no. 5, pp. 27–32, 2020.
- [22] R. Lima, B. Espinasse, and F. Freitas, "OntoILPER: an ontology- and inductive logic programming-based system to extract entities and relations from text," *Knowledge and Information Systems*, vol. 56, no. 1, pp. 223–255, 2018.