

## Research Article

# A Chinese Named Entity Recognition Model of Maintenance Records for Power Primary Equipment Based on Progressive Multitype Feature Fusion

Lanfei He,<sup>1</sup> Xuefei Zhang,<sup>1</sup> Zhiwei Li,<sup>1</sup> Peng Xiao,<sup>2</sup> Ziming Wei,<sup>2</sup> Xu Cheng,<sup>3</sup> and Shaocheng Qu<sup>ID</sup><sup>2</sup>

<sup>1</sup>Economic and Technical Research Institute of Hubei Electric Power Company of State Grid, Wuhan 430000, China

<sup>2</sup>Department of Electronics and Information Engineering, College of Physical Science and Technology, Central China Normal University, Wuhan 430000, China

<sup>3</sup>Wuhan Esmorning S&T Co., Ltd., Wuhan 430023, China

Correspondence should be addressed to Shaocheng Qu; [qushaocheng@mail.ccnu.edu.cn](mailto:qushaocheng@mail.ccnu.edu.cn)

Received 8 December 2021; Revised 30 December 2021; Accepted 11 January 2022; Published 7 February 2022

Academic Editor: Manman Yuan

Copyright © 2022 Lanfei He 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.

Presently, the State Grid Corporation of China has accumulated a large amount of maintenance records for power primary equipment. Unfortunately, most of these records are unstructured data which lead to difficulty analyze and utilize them. The emergence of natural language processing technology and deep learning methods provide a solution for unstructured text data. This paper proposes a progressive multitype feature fusion model to recognize Chinese named entity of unstructured maintenance records for power primary equipment. Firstly, the textual characteristics and word separation difficulties of maintenance records are analyzed, then 7 main entity categories of power technical terms from unstructured maintenance records are chosen, and 3452 maintenance records are labeled by these categories, which is so called EPE-MR training dataset. Secondly, the standard test reports, standard maintenance, and fault analysis reports for three types of power primary equipment (namely, main transformer, circuit breaker, and isolating switch) are employed as corpus to train character embedding in order to obtain certain words representation ability of maintenance records. After that, progressive multilevel radicals feature extraction module is designed to get detailed and fine semantic information in a hierarchical manner. Further, radicals feature representation and character embedding are concatenated and sent to BiLSTM module to extract contextual information in order to improve Chinese entity recognition ability. Moreover, CRF is introduced to handle the dependencies among prediction labels and to output the optimal prediction sequence, which can easily obtain structured data of maintenance records. Finally, comparative experiments on public MSRA dataset, China People's Daily corpus, and EPE-MR dataset are implemented, respectively, which show the effectiveness of the proposed method.

## 1. Introduction

With the development of artificial intelligence technology, extracting high-level structured semantic information such as entities, attributes, and relations from natural language to solve more advanced tasks in various industries is a prevailing research hot spot [1, 2]. In the field of power energy, the requirements of smart grid for correlation analysis and processing of effective information contained in big data of electric power are also continuously increasing. In particular,

it is coping with a large number of operation and maintenance records generated during the maintenance of various types of equipment. However, in the prior art, the management of equipment operation and maintenance records is limited to electronic storage and query of historical data. It is difficult to use these data to do some meaningful processing analysis of the actual work, which cannot meet the requirements of intelligent information management and delicacy management. The main challenge is how to recognize Chinese named entity of maintenance records in the form of unstructured

text. Supposing one can obtain Chinese named entity in uniform structured text of maintenance records of power equipment, it is possible to analyze and evaluate the key information of power operation and maintenance records (e.g., analyzing equipment lifetime). One of the most effective methods is to construct knowledge graphs.

Named entity recognition (NER) is a crucial foundational step in knowledge graph construction, as a large number of interest scenarios of involving information extraction, question answering systems, and machine translation [3–5]. Moreover, NER is widely used in the fields of chemistry [6], medicine [7, 8], and military affairs [9]. It aims to identify entities of a specific category from a piece of unstructured text. Initially, named entity recognition mainly employed rule-based and dictionary-based methods [10], which construct rule templates and dictionary libraries by linguists to perform rule matching on the text. However, there are some limitations. First, the construction of these rules depends on specific language, domain, and text style, while the construction process is complicated and time-consuming. Second, different rules need to be constructed for different fields with poor portability. Subsequently, scholars proposed statistical-based methods, including hidden Markov model (HMM), maximum entropy model (MEM), support vector machine (SVM), and conditional random field (CRF) [11–14]. Nevertheless, these models suffer from strong dependence on the corpus and slow convergence speed.

Recently, deep learning-based approaches have enabled end-to-end named entity recognition via neural networks that no longer rely on manually defined features [15, 16]. The most commonly used method is based on character embedding, and then the character embedding features are input into a long short-term memory (LSTM) with conditional random field (CRF) [8, 17]. However, this method cannot represent the multiple meanings of a word. The emergence of BERT alleviates this problem and is widely used in Chinese medicine [18, 19], Chinese idiom recommendation [20], and Chinese sentiment classification [21]. Another drawback is that LSTM cannot encode back-to-front information. In response to this problem, the proposal of bidirectional long short-term memory (BiLSTM) [22] provides an approach for the network to capture the two-way semantic dependence. Moreover, the radicals of Chinese characters represent different meanings that can provide useful information to the network, so this feature can be considered for inclusion in the network. An example is given in Figure 1. The meaning of character “消” is to disappear. It is broken down into two parts, “氵” and “肖.” “氵” represents water or river, and “肖” means small. These two radicals are merged to indicate the water is small, which means that the water is gone or almost gone.

Based on the above analysis, we propose a Chinese named entity recognition method based on progressive multitype feature fusion (PMTFF) to recognize unstructured maintenance records of power primary equipment. The main contributions are as follows:

- (1) We choose 7 main entity categories of power technical terms from unstructured maintenance records

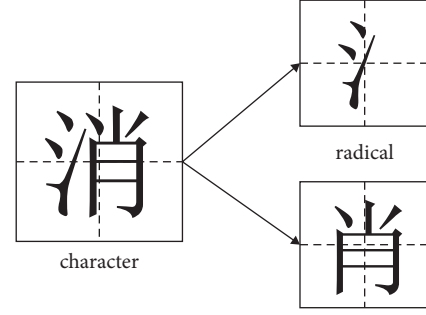


FIGURE 1: Decomposition of Chinese character.

and label 3452 unstructured maintenance records (EPE-MR dataset)

- (2) We employ the standard test reports, standard maintenance, and fault analysis reports of power primary equipment to train BERT model as the character embedding to obtain words representation ability
- (3) We propose a progressive multilevel radical feature extraction module (PML-RFE) to extract valuable semantic information

## 2. Dataset of Power Primary Equipment Maintenance Records

In the information database of power primary equipment of national grid, the basic information of equipment, maintenance records, and other data are stored. This information is of great significance to the management and maintenance of equipment, fault cause judgment, and fault differentiation analysis. However, the currently used manual extraction method is costly, inefficient, and not conducive to statistical analysis of the data. Therefore, we propose to use a deep learning-based approach for analysis, where a large number of training samples are indispensable. For this purpose, we construct the first EPE-MR dataset for Chinese named entity recognition of maintenance records of power primary equipment.

**2.1. Text Characteristics and Difficulties Analysis of Maintenance Records.** Compared with the general Chinese text, the state grid database of power primary equipment contains many types of information, and new entities are constantly emerging, such as damaged location of equipment, manufacture factory, and user unit. Meanwhile, the complexity of faults and repair methods varies from device to device, resulting in wide variations in the text length of different maintenance records. Statistics from the existing data show that the shortest overhaul records are 11 words, and the longest can be up to 354 words.

From the collected data and experiments, we conclude that the difficulty of word segmentation for maintenance records has the following three points:

- (i) The structure, integrity, and textual style of information are greatly different due to the different

writing habits of maintenance person. One of the more obvious cases is that some key information is abbreviated. For example, “主变电压器 (main voltage converter)” is abbreviated to “主变 (main transformer),” which creates a barrier to understanding maintenance records.

- (ii) The numerical information in the text relies on contextual feature for entity category judgment, such as voltage level “110kv,” equipment maintenance time “2019-09-0513:46:41,” and equipment model “电容器型号 (capacitor model) BAMH2.”
- (iii) For different maintenance records, the text describing the damaged parts and maintenance of equipment has the problems of tedious content, complex semantics, and unclear segmentation boundaries.

**2.2. Data Collection and Description.** We collect and label the power primary equipment maintenance records (EPE-MR) dataset from the enterprise resource planning (ERP) and power production management system (PMS) of Hubei State Grid Corporation. The labeling tool we used is Wizard Labeling Assistant [23], and six members of the group worked together to complete the annotation. In the task of named entity recognition, there are two commonly used identification systems: BIOES and BIO. We employ the BIO annotation approach to label the data. Concretely, B (Begin) represents the beginning of an entity, I (Intermediate) denotes the middle of an entity, and O (Other) represents the part that does not belong to any type of entity. In the constructed dataset, we define seven entities in power domain. Specifically, the first character is marked as “B (entity category),” the subsequent characters are marked as “I (entity category),” and other characters not related to the field are uniformly marked as O. We give an example in Table 1.

The entity classification of the existing maintenance records dataset is not complete enough, and the data volume is small, which brings challenges to the detection accuracy. To solve this issue, we construct the EPE-MR dataset, which contains seven entity categories of power terminology (i.e., VoltageLevel, EquipmentName, LineName, TransforSta, DamagePart, RepairCondition, and Time) and involves rich scenarios and a large amount of maintenance data. The entire dataset has a total of 3452 sentences, and we divide them into a training set and a test set at a ratio of 9 to 1. Figure 2 gives examples of sentences and entities in the EPE-MR dataset.

### 3. Proposed Method

The structure of the proposed model is shown in Figure 3. First, we get radical feature representation by the proposed progressive multilevel radical feature extraction module (PML-RFE). Then, the character embedding and radical feature representation are concatenated and sent to the BiLSTM module. Finally, the generated sequence is judged using a CRF module to obtain a globally optimal sequence.

#### 3.1. PMTFF Module

**3.1.1. Character Embedding.** In natural language processing, character embedding is usually employed to map a word to a low-dimensional dense semantic space, which can effectively solve the problem of text feature sparseness in traditional machine learning methods, so that similar words in the semantic space have closer distance [24, 25].

Pretrained character embeddings are beneficial to improve the accuracy of the model. Nevertheless, as far as we know, no scholars have given pretrained character embeddings for Chinese power primary equipment maintenance records. Thus, we employ our own EPE-MR dataset to train character embeddings. Currently, Word2Vec and GPT are more commonly used. However, these models have obvious problems with Chinese named power equipment maintenance records. Specifically, Word2Vec model is a static word embedding, which cannot represent multiple meanings of a word. GPT is a one-way language model, which cannot obtain the context of a word. The emergence of the BERT model alleviates the above problems. It enhances the semantic representation of word vectors by considering character-level and sentence-level relational features. Meanwhile, if the semantic knowledge is applied to the named entity recognition task in the power domain, it allows the model to better mine the feature information of power records and texts. Therefore, we utilize BERT to train Chinese power equipment maintenance records.

**3.1.2. Progressive Radical Feature Representation.** The development of Chinese characters consists of two main aspects. In terms of form, it gradually changes from graphics to strokes, pictographic to symbolic, and complexity to simplicity. In the principle of character creation, it has gone through a process from representational and ideographic to morpho-syntactic. Chinese characters are commonly construed as consisting of radicals and original basic parts. Radicals can reflect the intrinsic meaning of Chinese characters, which can provide valuable semantic knowledge for accurate model identification. Consequently, we propose to extract the characteristic representation of the radicals.

Chinese characters have different constructions and are square in shape, which is similar to image. We can exploit these similarities to obtain semantic knowledge using convolutional neural network (CNN). Here, we propose a progressive multilevel radical feature extraction module (PML-RFE). Figure 4 shows the overall schematic diagram of radical feature extraction, and the specific structure of PML-RFE is given in Figure 5.

Different hierarchical of characteristics contain different information. Low-level features possess rich spatial structure and deep-level information contains valuable semantic clues. Making full use of these two types of features can provide sufficient information for the network. However, if these two types of features are simply added or spliced by elements, it will cause redundancy of information and ignore the complementarity between different levels of features. Therefore, we propose the progressive attention mechanism

TABLE 1: Example of text labeling of maintenance records (钟226隔离开关发热跟踪测温 means clock 226 isolation switch fever tracking temperature measurement).

| Morpheme | Named entity annotation | Morpheme | Named entity annotation |
|----------|-------------------------|----------|-------------------------|
| 钟        | B-EquipmentName         | 关        | I-EquipmentName         |
| 2        | I-EquipmentName         | 发        | O                       |
| 2        | I-EquipmentName         | 热        | O                       |
| 6        | I-EquipmentName         | 跟        | B-RepairCondition       |
| 隔        | I-equipmentName         | 踪        | I-RepairCondition       |
| 离        | I-EquipmentName         | 测        | I-RepairCondition       |
| 开        | I-EquipmentName         | 温        | I-RepairCondition       |

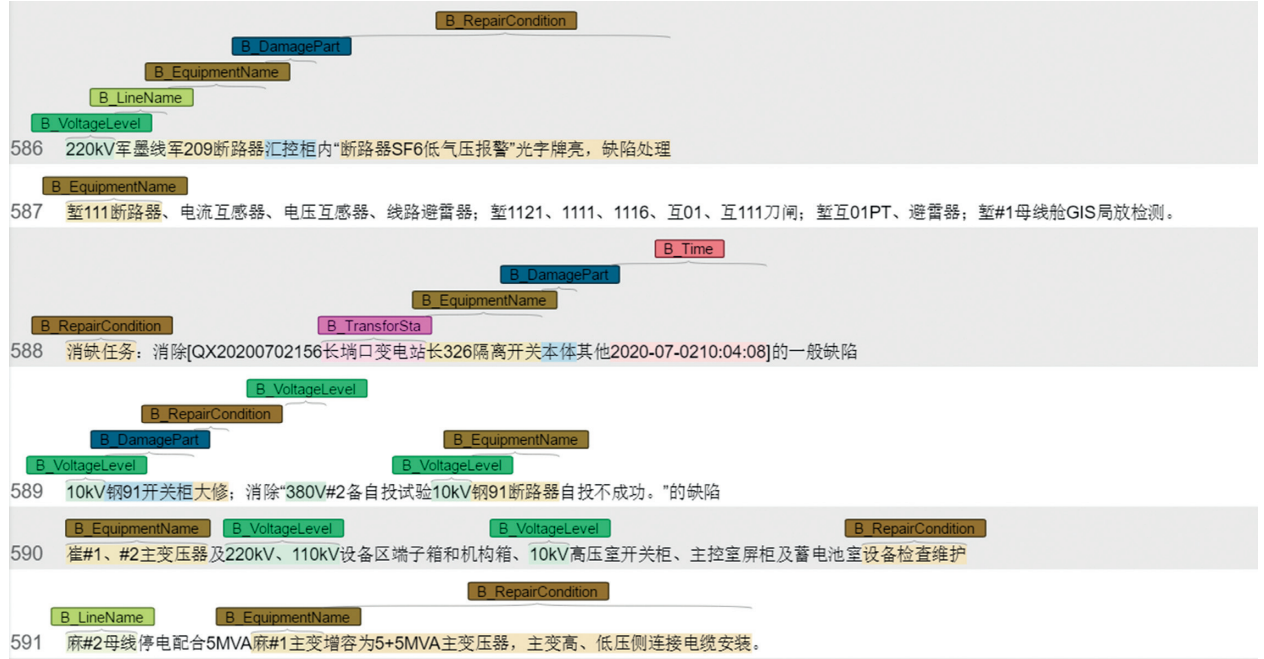


FIGURE 2: Example of EPE-MR dataset.

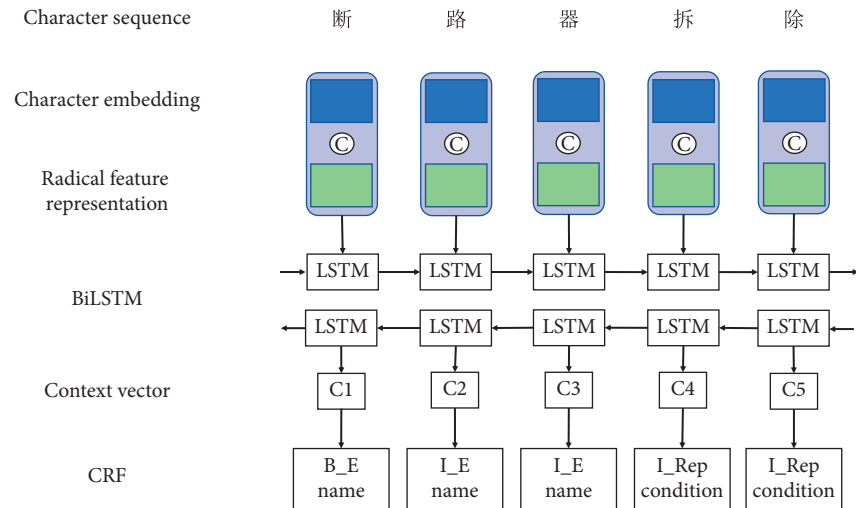


FIGURE 3: The framework of the proposed PMTFF model (温 means breaker removal).



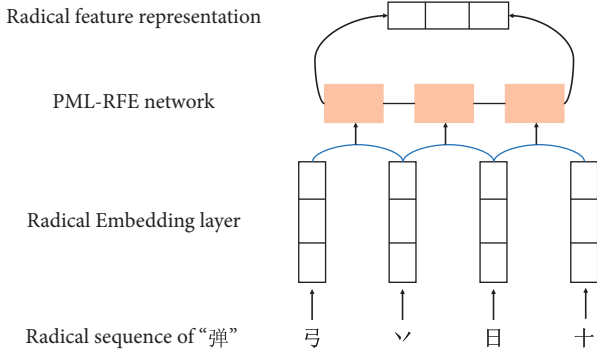


FIGURE 4: Architecture of radical feature extraction module.

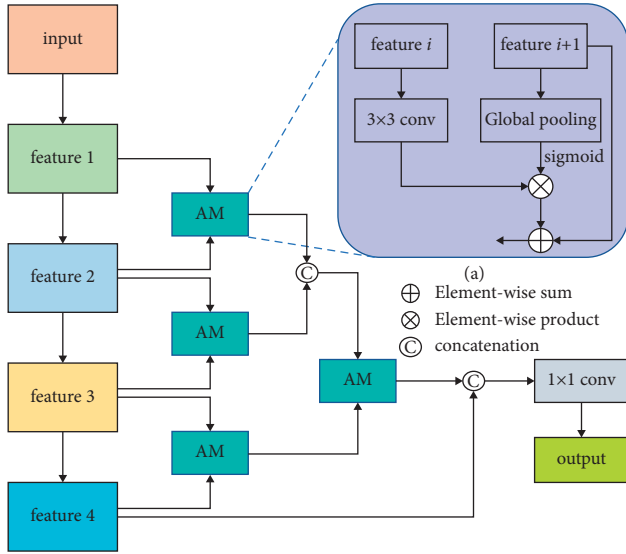


FIGURE 5: Structure of PML-RFE module.

to gradually acquire the connections between different levels of features, as shown in Figure 5.

First, four feature maps of different levels are obtained by CNN. Second, two adjacent feature maps are processed in turn to get enhanced features. Finally, this information is gradually merged instead of stitching them all together. More importantly, the enhanced features are obtained by the attention mechanism (AM) we designed, as shown in Figure 5(a). Specifically, the information obtained by the high-level features ( $F_r^{i+1}$ ) through global pooling and sigmoid activation function is used as weight to guide the low-level features ( $F_r^i$ ). Then, the weighted features ( $e_r^i \otimes e_r^{i+1}$ ) are merged with the high-level features and sent to the next stage. Equation (1) presents the expression of this process.

$$\begin{aligned} e_r^i &= f^{3 \times 3}(F_r^i), \\ e_r^{i+1} &= \sigma(GP(F_r^{i+1})), \\ F_r &= F_r^{i+1} \oplus (e_r^i \otimes e_r^{i+1}), \end{aligned} \quad (1)$$

while  $GP$  denotes the global pooling,  $\sigma$  is the sigmoid activation function,  $i = (1, 2, 3)$ , and  $f^{3 \times 3}(\cdot)$  represents  $3 \times 3$  convolution, and  $F_r$  is the new feature map.

**3.2. BiLSTM Module.** In NER task, recurrent neural network (RNN) is usually used to deal with sequence annotation problem. However, the traditional RNN cannot effectively solve the “long-range dependence” problem of sequence data [26, 27]. With the introduction of the LSTM neural network, this problem has been solved. LSTM is not only able to capture long-range information but also alleviates the gradient disappearance problem, so it is suitable for named entity recognition in the electric power field.

Figure 6 presents the basic structure of LSTM, including forgetting gate, input gate, and output gate. The long-term memory function is achieved by maintaining and updating the network state. The calculation procedure is shown in

$$\begin{aligned} i_t &= \sigma(W_i \cdot [h_{t-1}, x_t] + b_i), \\ f_t &= \sigma(W_f \cdot [h_{t-1}, x_t] + b_f), \\ O_t &= \sigma(W_o \cdot [h_{t-1}, x_t] + b_o), \\ C_t &= f_t * C_{t-1} + i_t * \tilde{C}_t, \\ h_t &= O_t * \tanh(C_t), \\ \tilde{C}_t &= \tanh(W_c \cdot [h_{t-1}, x_t] + b_c), \end{aligned} \quad (2)$$

where  $W$  is the weight matrix,  $b$  is the bias vector,  $\sigma$  is the activation function,  $\tilde{C}_t$  is the content to be added,  $C_t$  denotes the update state at time  $t$ ,  $i_t$ ,  $f_t$ , and  $O_t$  represent are the input gate, forget gate, and output gate at time  $t$ , respectively, and  $h_t$  is the output at time  $t$ .

The output prediction of LSTM is unidirectional and cannot process context information. However, the recognition of Chinese entities in power primary equipment maintenance records needs to be judged based on the status information before and after.

BiLSTM is a combination of forward LSTM and backward LSTM. Figure 7 gives the framework of BiLSTM. The execution steps are as follows: first, it calculates the input sequence in order and reverse order to obtain two different hidden layer representations. Then, the final hidden layer feature representation is obtained through vector splicing. By using BiLSTM, context relationships can be captured to improve the accuracy of named entity recognition.

**3.3. CRF Module.** Considering the correlation between adjacent labels, we add a condition random field (CRF) [28] inference layer after the BiLSTM network to obtain an optimal prediction sequence through the relationship between adjacent labels. The CRF algorithm steps are as follows:

Assuming that  $X = \{x_1, x_2, \dots, x_n\}$  is the complete expansion of the input sequence vector, the score of the predicted sequence  $Y = \{y_1, y_2, \dots, y_n\}$  is shown in

$$S(X, Y) = \sum_{i=0}^n A_{y_i, y_{i+1}} + \sum_{i=0}^n P_{i, y_i}, \quad (3)$$

while  $A$  is the transfer score matrix,  $A_{ij}$  represents the transfer score from label  $i$  to label  $j$ , matrix  $P$  is the output of

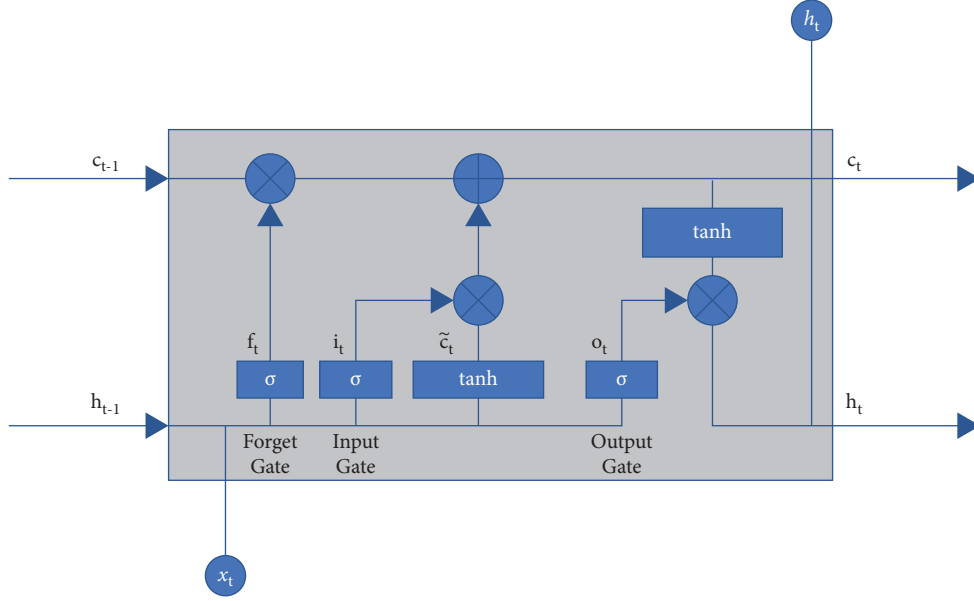


FIGURE 6: The structure of LSTM.

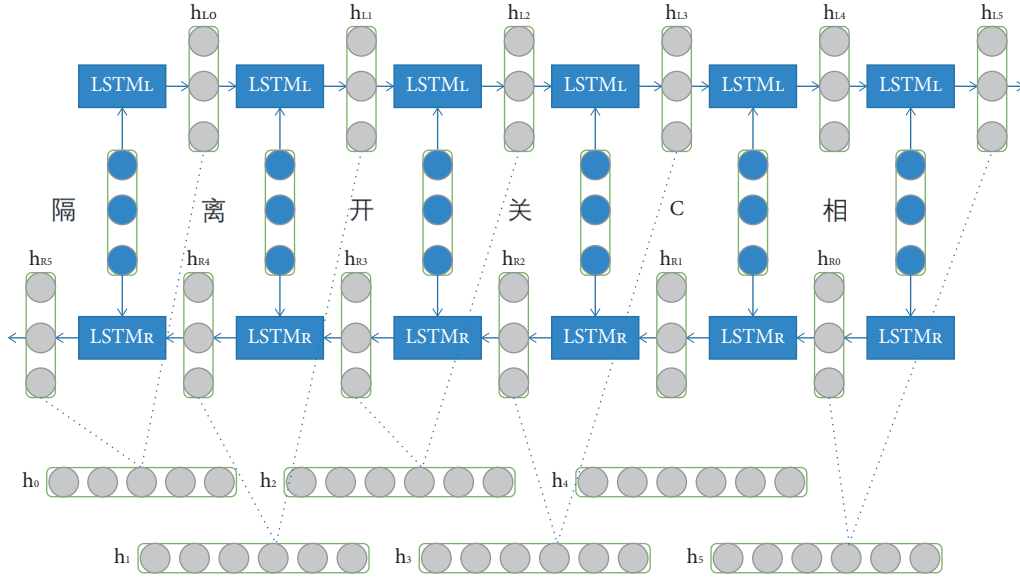


FIGURE 7: The structure of BiLSTM.

the BiLSTM layer,  $P_{ij}$  represents the output score of the  $i$ -th word under the  $j$ -th label,  $n$  represents the sequence length, and  $k$  represents the number of labels. The probability of predicting sequence  $Y$  is calculated by

$$P(Y|X) = \frac{e^{S(X,Y)}}{\sum_{\tilde{Y} \in Y_x} e^{S(X,\tilde{Y})}}, \quad (4)$$

while  $\tilde{Y}$  represents the real marker sequence and  $Y_x$  is all possible marker sequences. During the training process, maximize the likelihood probability of the correct label sequence, as shown in

$$\log a^{P(Y|X)} = S(X, Y) - \sum_{\tilde{Y} \in Y_x} S(X, \tilde{Y}). \quad (5)$$

The output sequence with the maximum score after decoding is shown in

$$Y^* = \arg \max_{\tilde{Y} \in Y_x} S(X, \tilde{Y}). \quad (6)$$

## 4. Experiments

**4.1. Datasets and Evaluation Index.** We perform experiments on our own constructed power primary equipment

maintenance records (EPE-MR) dataset, MSRA [29] dataset, and China People’s Daily corpus.

- (1) EPE-MR is a dataset we made for maintenance records of power primary equipment. There are 3452 sentences and 7 types of entities, which are VoltageLevel, EquipmentName, LineName, TransforSta, DamagePart, RepairCondition, and Time. As far as we know, this is the first Chinese entity recognition dataset for the subject.
- (2) MSRA dataset is a public Chinese entity recognition dataset. Specifically, there are 46364 and 4365 sentences in the training set and test set, respectively. The entity types included are person, organization, and location.
- (3) China People’s Daily corpus contains the same types of entities as MSRA. The dataset provides sufficient data information, with 17573 and 1718 sentences in the training and test sets, respectively.

The evaluation indexes used in the experiment are precision ( $P$ ), recall ( $R$ ), and  $F1$ . The calculation methods of each evaluation index are as follows:

$$\begin{aligned}
 R &= \frac{TP}{TP + FN}, \\
 P &= \frac{TP}{TP + FP}, \\
 F1 &= \frac{2 \times R \times P}{R + P},
 \end{aligned} \tag{7}$$

while  $TP$  is the number of entities that are correctly classified,  $FN$  is the number of entities that are actually related entities but not recognized by the model, and  $FP$  is the number of entities that are actually unrelated entities but the model judges them to be entities.

**4.2. Ablation Study.** We implement the proposed model with PyTorch in Python 3.6.90% of the data are used as training set and the rest as test set. In the training process, the optimizer employs Adam with a learning rate of 0.001. Meanwhile, LSTM\_dim is set to 300, batch\_size to 32, and max\_seq\_len to 300. To prevent overfitting, Dropout is set to 0.5.

We perform ablation analysis on the EPE-MR dataset. Table 2 gives the performance analysis of different modules. The variations of  $F1$  values for different modules on the test set are presented in Figure 8. Specifically, the  $F1$  of our approach is 88.3%. From the perspective of comprehensive evaluation index  $F1$ , the proposed method is 14.2%, 9.9%, 12.1%, 8.8%, 7.6%, and 2.1% higher than other combinations. There are two main reasons for the success of proposed approach: (i) the proposed method enhances the semantic representation by fully considering the character-level, word-level, and radical-level relationship features; (ii) the two-way structure helps the network to obtain contextual information.

TABLE 2: Performance comparison of different modules on EPE-MR dataset (CE: character embedding; RFR: radical feature representation).

| Modules                                   | Metrics      |              |              |
|---|--------------|--------------|--------------|
|   | $P$          | $R$          | $F1$         |
| LSTM                                      | 0.815        | 0.707        | 0.741        |
| BiLSTM                                    | 0.811        | 0.776        | 0.784        |
| CE  | 0.774        | 0.743        | 0.762        |
| CE-CRF                                    | 0.828        | 0.769        | 0.795        |
| BiLSTM-CRF                                | 0.834        | 0.805        | 0.807        |
| CE-BiLSTM-CRF                             | 0.869        | 0.866        | 0.862        |
| CE + RFR-BiLSTM-CRF (the proposed method) | <b>0.892</b> | <b>0.886</b> | <b>0.883</b> |

Bold values show the best performance.

For a more visual observation, the entity recognition results are given in Figure 9. Specifically, the left side of the figure shows the entity label types, and the right side presents the corresponding classification results. It can be seen that the proposed method has a good word separation effect on professional vocabulary in the electric power domain.

**4.3. Comparison with State-of-the-Arts Approaches.** To evaluate the proposed method more comprehensively, we compare it with Dong et al. [30], Lattice-LSTM-CRF [31], Cao et al. [32], WC-LSTM [33], and CNN-BiLSTM-CRF [34] on the MSRA dataset, as shown in Table 3. Specifically, Lattice-LSTM-CRF improved the traditional LSTM unit into a grid LSTM and then explicitly used word and word order information to avoid the transmission of word segmentation errors. WC-LSTM employed word information to enhance semantic information and reduce the impact of word segmentation errors. The above improved models always stay in the extraction of character and word features and cannot characterize both meanings simultaneously. In contrast, the proposed approach can solve this problem very well. It can learn not only the phrase-level information representation but also the rich semantic information features.

As can be seen from Table 3, the  $F1$  value of the proposed method reaches 96.63%, which is 5.68%, 3.45%, 5.99%, 2.89%, and 5.54% higher than that of Dong et al. [30], Lattice-LSTM-CRF [31], Cao et al. [32], WC-LSTM [33], and CNN-BiLSTM-CRF [34], respectively. The bar graphs of each quantitative indicator for the different methods are given in Figure 10. On the whole, the proposed approach has certain advantages.

Table 4 gives a quantitative comparison of Collobert et al. [35], Chiu and Nichols [36], Shen et al. [37], Lample et al. [38], and the proposed method on China People’s Daily corpus. The  $F1$  value of the proposed method reaches 96.54%. Figure 11 depicts the bar graph of quantitative analysis metrics for different models. From the comprehensive view of the three evaluation metrics, the proposed approach shows the best performance. Moreover, during the experiment, we found that the performance of the proposed model will decrease when it encounters typos.

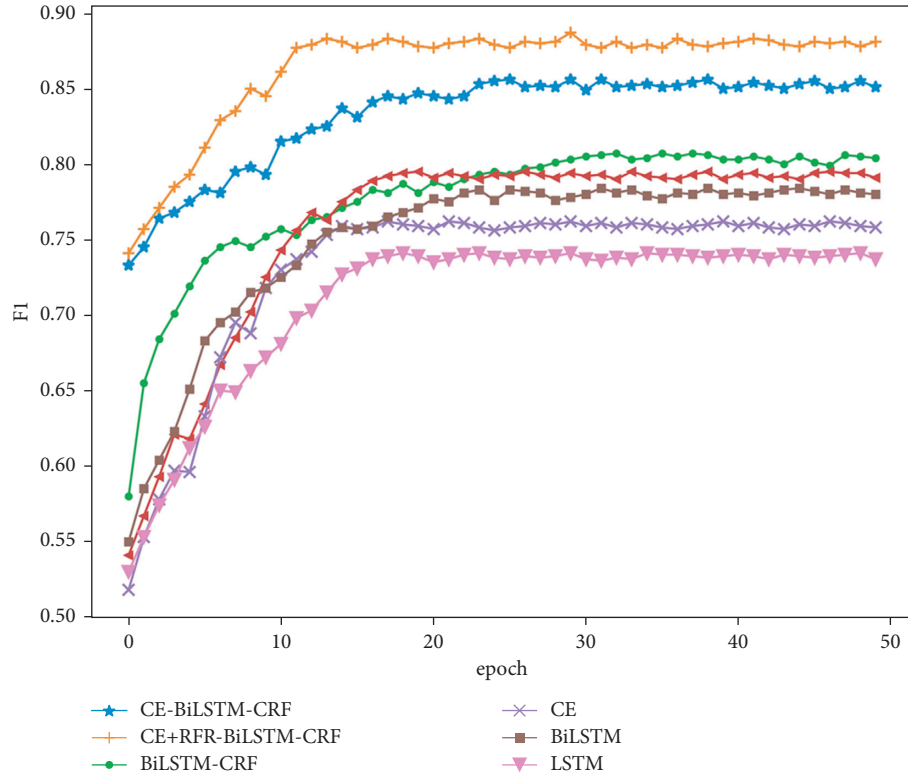
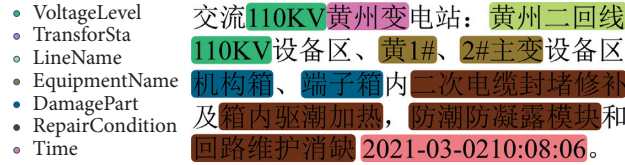
FIGURE 8: The variation of  $F1$  values for different modules on the test set.

FIGURE 9: Entity recognition results.

TABLE 3: Performance comparison of different methods on MSRA dataset.

| Methods                | Metrics      |              |              |
|------------------------|--------------|--------------|--------------|
|                        | $P$ (%)      | $R$ (%)      | $F1$ (%)     |
| Dong et al. [30]       | 91.28        | 90.62        | 90.95        |
| Lattice-LSTM-CRF [31]  | 93.57        | 92.79        | 93.18        |
| Cao et al. [32]        | 91.73        | 89.58        | 90.64        |
| WC-LSTM [33]           | 94.58        | 92.91        | 93.74        |
| CNN-BiLSTM-CRF [34]    | 91.63        | 90.56        | 91.09        |
| <b>Proposed method</b> | <b>97.85</b> | <b>95.74</b> | <b>96.63</b> |

Bold values show the best performance.



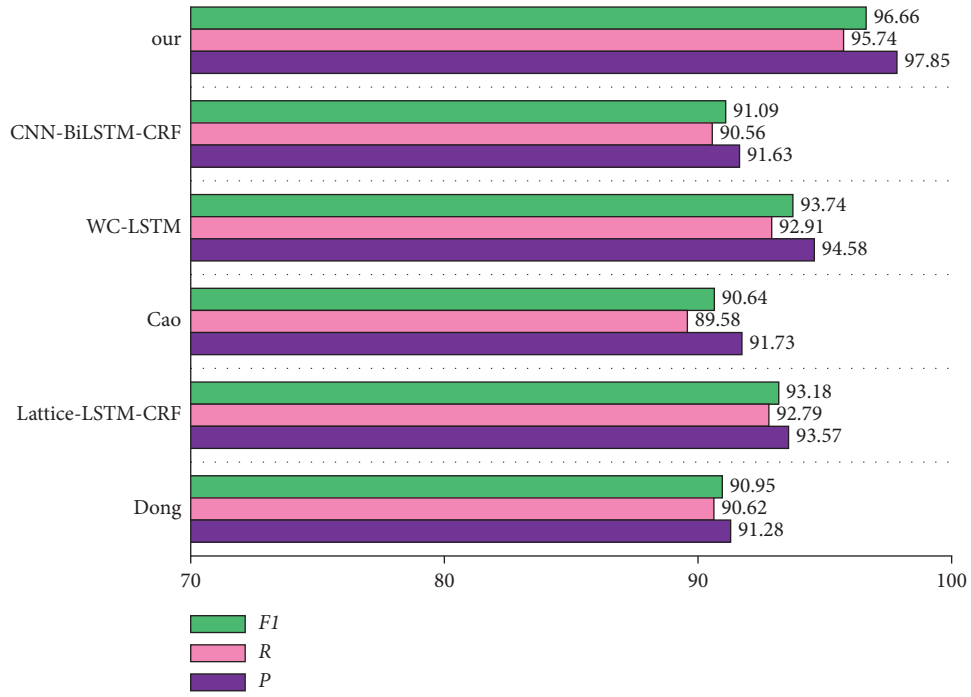
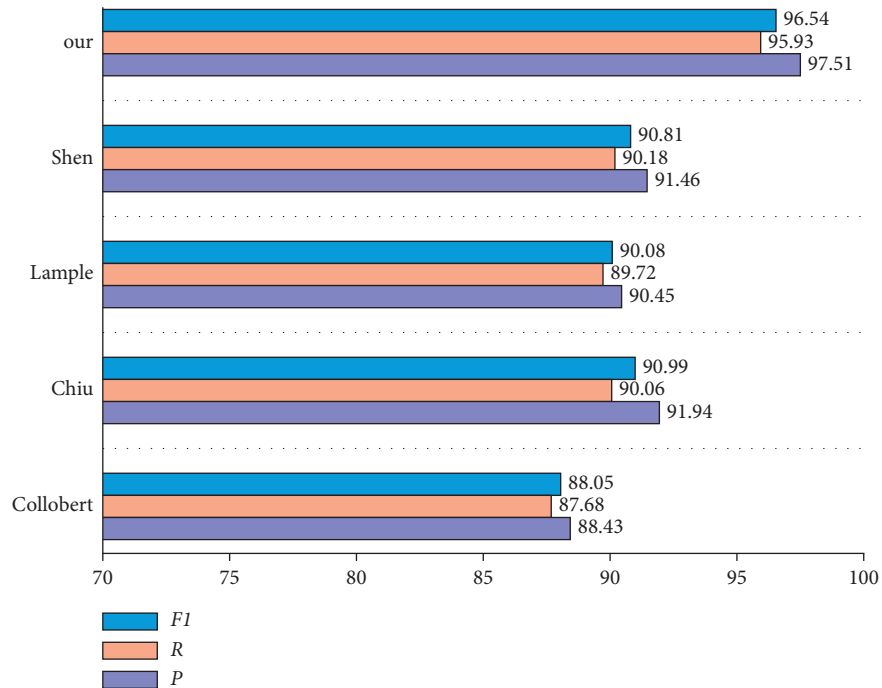
FIGURE 10:  $F1$ ,  $R$ , and  $P$  comparison of different methods on MSRA dataset.

TABLE 4: Performance comparison of different methods on China People's Daily corpus.

| Methods               | Metrics      |              |              |
|-----------------------|--------------|--------------|--------------|
|                       | $P$ (%)      | $R$ (%)      | $F1$ (%)     |
| Collobert et al. [35] | 88.43        | 87.68        | 88.05        |
| Chiu and Nichols [36] | 91.94        | 90.06        | 90.99        |
| Shen et al. [37]      | 90.45        | 89.72        | 90.08        |
| Lample et al. [38]    | 91.46        | 90.18        | 90.81        |
| The proposed method   | <b>97.51</b> | <b>95.93</b> | <b>96.54</b> |

Bold values show the best performance.

FIGURE 11:  $F1$ ,  $R$ , and  $P$  comparison of different methods on China People's Daily corpus.

## 5. Conclusion

In this paper, we propose a Chinese named entity recognition method based on progressive multitype feature fusion for maintenance records of power primary equipment. Firstly, a dataset for Chinese named entity recognition of power primary equipment maintenance records is collected and labeled. Then, the contextualized word vector is obtained using the BERT preprocessing model. Thirdly, radical feature representations are extracted by the PML-RFE module, and these features are concatenate with character embeddings and sent to the BiLSTM module. Finally, CRF is combined to improve the recognition effect. The experimental results on EPE-MR, MSRA, and China People's Daily datasets show that the proposed method achieves superior results compared with other existing methods.

## Data Availability

Our research is still in the research stage, so we cannot disclose relevant information of the data for the time being.

## Conflicts of Interest

There are no conflicts of interest from the authors.

## References

- [1] M. Yin, C. Mou, K. Xiong, and J. Ren, "Chinese clinical named entity recognition with radical-level feature and self-attention mechanism," *Journal of Biomedical Informatics*, vol. 98, Article ID 103289, 2019.
- [2] G. Jin and Z. Yu, "A Korean named entity recognition method using Bi-LSTM-CRF and masked self-attention," *Computer Speech & Language*, vol. 65, Article ID 101134, 2021.
- [3] M. Miwa and M. Bansal, "End-to-end relation extraction using LSTMs on sequences and tree structures," 2016, <http://arxiv.org/abs/1601.00770>.
- [4] Z. Yang, N. Garcia, C. Chu, M. Otani, Y. Nakashima, and H. Takemura, "A comparative study of language transformers for video question answering," *Neurocomputing*, vol. 445, pp. 121–133, 2021.
- [5] M. Maimaiti, Y. Liu, H. Luan, and M. Sun, "Enriching the transfer learning with pre-trained lexicon embedding for low-resource neural machine translation," *Tsinghua Science and Technology*, vol. 27, no. 1, pp. 150–163, 2022.
- [6] L. Luo, Z. Yang, P. Yang et al., "An attention-based BiLSTM-CRF approach to document-level chemical named entity recognition," *Bioinformatics*, vol. 34, no. 8, pp. 1381–1388, 2018.
- [7] Y. Li, X. Wang, L. Hui et al., "Chinese clinical named entity recognition in electronic medical records: development of a lattice long short-term memory model with contextualized character representations," *JMIR Medical Informatics*, vol. 8, no. 9, Article ID e19848, 2020.
- [8] X. Li, H. Zhang, and X.-H. Zhou, "Chinese clinical named entity recognition with variant neural structures based on BERT methods," *Journal of Biomedical Informatics*, vol. 107, Article ID 103422, 2020.
- [9] M. Zhou and J. Liu, "A memetic algorithm for enhancing the robustness of scale-free networks against malicious attacks," *Physica A: Statistical Mechanics and Its Applications*, vol. 410, pp. 131–143, 2014.
- [10] N. Zhang, G. Xu, Z. Zhang, and F. Li, "MIFM: multi-granularity information fusion model for Chinese named entity recognition," *IEEE Access*, vol. 7, pp. 181648–181655, 2019.
- [11] G. Zhou and J. Su, "Named entity recognition using an HMM-based chunk tagger," in *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, Philadelphia Pennsylvania, July 2002.
- [12] J. Lei, B. Tang, X. Lu, K. Gao, M. Jiang, and H. Xu, "A comprehensive study of named entity recognition in Chinese clinical text," *Journal of the American Medical Informatics Association*, vol. 21, no. 5, pp. 808–814, 2014.
- [13] Y. Li, K. Bontcheva, and H. Cunningham, "SVM based learning system for information extraction," in *Proceedings of the International Workshop on Deterministic and Statistical Methods in Machine Learning*, pp. 319–339, Springer, Sheffield, UK, September 2004.
- [14] S. N. V. P. Mitra and S. K. Ghosh, "Conditional random field based named entity recognition in geological text," *International Journal of Computer Application*, vol. 1, no. 3, pp. 143–147, 2010.
- [15] G. Aras, D. Makaroğlu, S. Demir, and A. Cakir, "An evaluation of recent neural sequence tagging models in Turkish named entity recognition," *Expert Systems with Applications*, vol. 182, Article ID 115049, 2021.
- [16] D. Jiang, H. Ren, Y. Cai, J. Xu, Y. Liu, and H.-f. Leung, "Candidate region aware nested named entity recognition," *Neural Networks*, vol. 142, pp. 340–350, 2021.
- [17] H. Wu, J. Ji, H. Tian et al., "Chinese-named entity recognition from adverse drug event records: radical embedding-combined dynamic embedding-based BERT in a bidirectional long short-term conditional random field (Bi-LSTM-CRF) model," *JMIR Medical Informatics*, vol. 9, no. 12, p. e26407, 2021.
- [18] Q. Qin, S. Zhao, and C. Liu, "A BERT-BiGRU-CRF model for entity recognition of Chinese electronic medical records," *Complexity*, vol. 2021, Article ID 6631837, 11 pages, 2021.
- [19] J. Wang, G. Zhang, W. Wang, K. Zhang, and Y. Sheng, "Cloud-based intelligent self-diagnosis and department recommendation service using Chinese medical BERT," *Journal of Cloud Computing*, vol. 10, no. 1, 2021.
- [20] M. Tan, J. Jiang, and B. T. Dai, "A BERT-based two-stage model for Chinese chengyu recommendation," *ACM Transactions on Asian and Low-Resource Language Information Processing*, vol. 20, no. 6, pp. 1–18, 2021.
- [21] J. Gao, "Chinese sentiment classification model based on pre-trained BERT," in *Proceedings of the 2021 2nd International Conference on Computers, Information Processing and Advanced Education*, pp. 1296–1300, Ottawa, Canada, May 2021.
- [22] A. Graves and J. Schmidhuber, "Framewise phoneme classification with bidirectional LSTM and other neural network architectures," *Neural Networks: The Official Journal of the International Neural Network Society*, vol. 18, no. 5-6, pp. 602–610, 2005.
- [23] <https://www.jinglingbiaozhu.com>.
- [24] M. Pota, M. Ventura, H. Fujita, and M. Esposito, "Multilingual evaluation of pre-processing for BERT-based sentiment analysis of tweets," *Expert Systems with Applications*, vol. 181, Article ID 115119, 2021.
- [25] A. Zhao and Y. Yu, "Knowledge-enabled BERT for aspect-based sentiment analysis," *Knowledge-Based Systems*, vol. 227, Article ID 107220, 2021.
- [26] L. Shan, Y. Liu, M. Tang, M. Yang, and X. Bai, "CNN-BiLSTM hybrid neural networks with attention mechanism for well log

- prediction,” *Journal of Petroleum Science and Engineering*, vol. 205, Article ID 108838, 2021.
- [27] N. Qun, H. Yan, X.-P. Qiu, and X.-J. Huang, “Chinese word segmentation via BiLSTM + semi-CRF with relay node,” *Journal of Computer Science and Technology*, vol. 35, no. 5, pp. 1115–1126, 2020.
  - [28] J. Lafferty, A. McCallum, and F. Pereira, “Conditional random fields: probabilistic models for segmenting and labeling sequence data,” in *Proceedings of the 18th International Conference on Machine Learning*, pp. 1–8, Williamstown, MA, USA, July 2001.
  - [29] G. Levow, “The third international Chinese language processing bakeoff: word segmentation and named entity recognition,” in *Proceedings of the Fifth Workshop on Chinese Language Processing*, pp. 108–117, Sydney, Australia, July 2006.
  - [30] C. Dong, J. Zhang, C. Zong, M. Hattori, and H. Di, “Character-based LSTM-CRF with radical-level features for Chinese named entity recognition,” *Natural Language Understanding and Intelligent Applications*, Springer, Berlin, Germany, pp. 239–250, 2016.
  - [31] Y. Zhang and J. Yang, “Chinese NER using lattice LSTM,” 2018, <https://arxiv.org/abs/1805.02023>.
  - [32] P. Cao, C. Yu, L. Kang, Z. Jun, and S. Liu, “Adversarial transfer learning for Chinese named entity recognition with self-attention mechanism,” in *Proceedings of the Empirical Methods in Natural Language Processing*, pp. 182–192, Brussels, Belgium, October 2018.
  - [33] W. Liu, X. Tongge, X. Qinghua, and S. Jiayu, “An encoding strategy based word-character LSTM for Chinese NER,” in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pp. 2379–2389, Minneapolis, Minnesota, June 2019.
  - [34] Y. Jia and X. Xu, “Chinese named entity recognition based on CNN-BiLSTM-CRF,” in *Proceedings of the 9th International Conference on Software Engineering and Service Science*, pp. 161–175, Springer Singapore, Beijing, China, November 2018.
  - [35] R. Collobert, J. Weston, L. Bottou, M. Karlen, K. Kavukcuoglu, and P. Kuksa, “Natural language processing (almost) from scratch,” *Journal of Machine Learning Research*, vol. 12, no. 12, pp. 2493–2537, 2011.
  - [36] J. P. C. Chiu and E. Nichols, “Named entity recognition with bidirectional LSTM-CNNs,” *Transactions of the Association for Computational Linguistics*, vol. 4, pp. 357–370, 2016.
  - [37] Y. Shen, H. Yun, and Z. C. Lipton, “Deep active learning for named entity recognition,” 2017, <https://arxiv.org/abs/1707.05928v3>.
  - [38] G. Lample, M. Ballesteros, S. Subramanian, K. Kawakami, and C. Dyer, “Neural architectures for named entity recognition,” in *Proceedings of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, San Diego, CA, USA, June 2016.