DdERT: Research on Named Entity Recognition for Mine Hoist Using a Chinese BERT Model

Xiaochao Dang¹,* , Li Wang², Xiaohui Dong¹,², Fenfang Li² and Han Deng²

¹ College of Computer Science and Engineering, Northwest Normal University, Lanzhou 730070, China; dxh@nwnu.edu.cn
² Gansu Province Internet of Things Engineering Research Center, Lanzhou 730070, China;
2021222195@nwnu.edu.cn (L.W.); lifenfang@nwnu.edu.cn (F.L.); 2021222276@nwnu.edu.cn (H.D.)
* Correspondence: dangxc@nwnu.edu.cn

Abstract: This study aims to solve the problem of named entity recognition of complex mechanical equipment faults, especially the problems of many professional terms, long sentences, fuzzy entity boundaries, entity nesting, and abbreviation ambiguity, in mine hoist fault text. Therefore, this study proposes a named entity recognition method based on domain dictionary embedding. The method first uses the fault domain knowledge of the mine hoist to construct a domain-specialized dictionary and generate a word vector of characteristic words. Secondly, the BERT pre-trained language model is used to obtain dynamic word vectors, and a dictionary adapter is loaded to obtain contextual domain lexical features to improve recognition accuracy. Finally, the conditional random field (CRF) is the model classifier to output the annotation sequence with the highest score. The experimental results show that this model achieves better than several baseline models and effectively improves the accuracy of fault named entity identification for mine hoists. The innovation of this study is the combination of domain dictionary embedding and a BERT pre-trained language model, which improves the accuracy and robustness of named entity recognition. Therefore, the results of this study have essential research significance for improving the accuracy of fault named entity identification of mine hoists and the construction of fault knowledge maps.

Keywords: knowledge extraction; mine hoist fault; named entity recognition; BERT; construction of fault knowledge maps; deep learning; information extraction; fault diagnosis; mining sector

1. Introduction

Mine hoists are the leading transportation equipment in mining engineering. Once failure occurs, it will cause a significant loss to enterprises’ production and even endanger underground operators’ safety. At present, the operation and maintenance of mine lifting equipment mainly rely on the subjective experience of maintenance personnel, and it is necessary to repeatedly consult a large amount of fault maintenance information, such as operation manuals, maintenance methods, and safety procedures [1–3]. Therefore, a more efficient maintenance method is needed to improve lifting equipment operation efficiency and safety.

As a means of knowledge expression based on artificial intelligence technology, knowledge graph has shown its application potential in many fields, including information extraction, maintenance scheme generation, auxiliary fault location, and other equipment operation and maintenance fields [4–7]. Knowledge graphs can deal with complex maintenance knowledge efficiently and facilitate the sharing of manual experience. Therefore, using knowledge graph technology to enhance equipment operation and maintenance efficiency, and safety has been regarded as a disruptive maintenance strategy. However, in constructing a knowledge graph, named entity recognition [8] (NER) is a crucial step, and its recognition effect directly impacts the knowledge graph’s quality. Entity identification
tasks can mine valid information in fault text, transform unstructured data into structured data, and effectively improve the efficiency of knowledge storage and query. Owing to the fuzzy entity boundary, many terms, insufficient context information, entity nesting, and abbreviated ambiguous entity in mine hoist fault text, the traditional entity recognition method needs to be more efficient in training and challenging to improve the recognition effect. How to accurately and efficiently identify, mine, and analyze all kinds of entity information in the equipment fault text plays a crucial role in the intelligent decision of equipment maintenance.

In summary, how to accurately extract fault information, such as fault causes, fault phenomena, and maintenance measures, from a large number of redundant unstructured texts and solve the problem of entity nesting and abbreviation ambiguity in the text is the core problem and challenge in mine hoist fault knowledge graph construction. Given the above problems, this study proposes a mine hoist fault recognition model (domain dictionary embedding model based on BERT, DdERT). It is verified on the fault data set of a self-built mine hoist. The experimental results show that the proposed model can effectively improve the effect of entity recognition.

Aiming at the above problems, based on the BERT pre-trained language model, this study proposes a mine hoist fault entity recognition model, DdERT. The main contributions of this work are as follows:

- Constructing a data set in the field of mine hoist faults: We conducted an innovative study to explore the dataset of mine hoist faults in detail, which has not been fully filled in the current field. To date, there is no dataset that covers both entity recognition and fault diagnosis. This study closely combines internal data from mining companies and the publicly available literature, and carefully constructs a high-quality dataset of mine hoist fault data through manual annotation. This dataset not only provides detailed references, but also greatly promotes research and development in this field. Although this dataset cannot be publicly released, it provides strong support for relevant research and is widely used in academic research;
- Construct the DdERT model: To embed the elevator fault dictionary BERT model, a modified coding layer fusion algorithm is combined with a dictionary for each input character for training. At the same time, conditional random field (CRF) [9] was used as the model classifier to alleviate the imbalance problem between samples in the mine lifting field and further improve the effect of the DdERT model. Compared with other models, the F1 value of the DdERT model was greatly improved.

2. Related Work

In the early research in the field of NER, the mainstream method was the rule-based method [10,11]. However, the limitation of this method is that the rules need to be created manually, and because named entity types are various and constantly evolving, the required domain knowledge is also constantly changing. Hence, the maintenance and update of rules take a lot of time and effort. To overcome these limitations, rule-based methods are being combined with machine learning models to improve NER performance [12,13].

In 2018, Devlin et al. proposed a revolutionary pre-trained language model, BERT [14], which achieved breakthrough results on SQuAD1.1, the top evaluation dataset for machine reading comprehension. It was a milestone in natural language processing (NLP). The BERT model has subsequently been widely used in NER tasks. For example, Souza et al. [15] proposed a BERT-CRF model that combined BERT with CRF. Li et al. [16] applied the BERT model for pre-training on unlabeled Chinese clinical electronic medical record text, integrated dictionary features into the model, and used Chinese character root features to enhance the performance of NER tasks further. On this basis, Wu et al. [17] proposed a method of using a bi-directional long–short-term memory (BiLSTM) neural network [18] to extract partial radical features by combining RoBERTa and character root features.
features. Furthermore, a model combining these features with the medical feature vectors learned by RoBERTa further enhanced the performance of NER tasks.

The BERT model also shows advantages in named entity recognition in the industrial field. Liang et al. [19] combined BERT with self-training methods and optimized the model on many automatically annotated corpora to adapt to NER tasks in low-resource environments. Liu et al. [20] obtained character-level features through the character vector representation layer of BERT and introduced an improved attention mechanism to capture local features to improve the entity recognition performance of the power fault corpus. Baigang et al. [21] combined BERT with transfer learning and proved the effectiveness of this model on an aviation maintenance dataset.

It is worth noting that, in recent years, incorporating lexical information into the model for lexical augmentation has proven to be an effective way to improve the performance of Chinese NER. For example, Liu et al. [22] proposed a model called LEBERT, which effectively performs word augmentation by injecting lexical features into a specific layer of BERT. Li et al. [23] proposed a multi-modal domain knowledge graph construction method based on LEBERT to solve the problem of large and scattered knowledge systems in computer science. Wu et al. [24] proposed a LEBERT-BCF model by combining an external dictionary with BERT features and adopting an adversarial training strategy, which solved multiple problems in the entity recognition of Chinese electronic medical records.

However, the current study of mine hoists on NER is still an unexplored area. Most of them combine mine hoists with deep learning technology and data-driven methods to achieve fault prediction and diagnosis. Li et al. [25] proposed the motor fault diagnosis method based on CNN-BiLSTM to identify the input MFCC feature vector, and the accuracy reached 92.78% through comparative experiments, which effectively improved the accuracy of the fault diagnosis method. Ruan et al. [26] used data-driven methods to make twin models use knowledge to make auxiliary decisions and to realize the global visualization of the mine hoist system, rapid fault localization, and troubleshooting. Guo et al. [27] used a mine hoist early warning model based on LSTM-Adam to predict the parameter change trend and used the prediction residual analysis results to accurately predict the hoist fault. The experimental results show that the fault early warning model based on LSTM-Adam could warn of the hoist fault phenomenon in time. Cao [28] used cutting-edge technologies, such as high-precision sensing technology, AI video analysis technology, and industrial Internet technology, to design an intelligent remote monitoring system for mine hoists. Real-time monitoring of system operation status parameters, accurate identification, and early warning of various types of faults significantly shortened the fault disposal time.

The above studies lay the foundation and provide solid theoretical support for the work of this study. However, as a specific domain, mine hoists’ fault texts usually contain rich technical terms and special symbols, characterized by lengthy statements, numerous entity categories with unclear boundaries, and ambiguity of entity nesting and abbreviation. These characteristics make mine hoist fault texts significantly different from general domain texts. Take “improper gear meshing” as an example; this phrase is composed of “gear” and “improper meshing”, which may be mistakenly identified as the entity type of a fault location and fault phenomenon in traditional entity recognition methods, while ignoring that it is a cause of fault phenomena. In addition, traditional Chinese entity recognition methods usually require text segmentation. However, word segmentation only depends on the character level that may lead to the loss of semantic information, and the errors of the word segmentation model itself may accumulate and affect the entity recognition model, thereby reducing the accuracy of entity recognition. At the same time, directly transferring entity recognition models from general domains to specific domains, such as mine hoist fault texts, usually does not produce desirable recognition results. This is mainly because these models often fail to effectively identify the professional terms and their boundaries in the domain, resulting in the recognition of entities that cannot accurately express their accurate semantic information. Therefore, for entity recognition of mine hoist fault texts,
it is necessary to develop more refined and adaptable models to accurately capture and recognize the professional terms and entities in the field to improve the accuracy and reliability of entity recognition.

3. Model

The DdERT model combining the Chinese BERT model and the lexicon adapter was used for the named entity recognition task. As illustrated in Figure 1, the entire architecture had three levels:

1. Input layer: The input layer processes the input text. In this layer, we employ the BERT model, which can map the individual words and terms in the input text to their corresponding vector representations. These vector representations, including character-level and word-level vectors, are combined as the input to the model.

2. Dictionary fusion encoding layer: The core purpose of this layer is to fuse the character-level feature vectors and the word-level feature vectors. To achieve this, the lexicon fusion encoding layer interjects a lexicon adapter between the two transformer encoders. By scanning the dictionary adapter, the model can match the character-level feature vectors with the corresponding word-level feature vectors and learn how to fuse the two feature vectors. This fusion helps the model to identify entity boundaries more accurately and improve its encoding accuracy of semantic information.

3. Decoding Layer: The decoding layer further processes the character-level and word-level feature vectors based on the dictionary fusion encoding layer. After the composite operation of multiple transformer layers and dictionary adapters, the model can learn the fusion weights of these feature vectors. Then, CRF is used to fine-tune the entity recognition results of the model to output the globally optimal annotation sequence.

Figure 1. An overview of the DdERT model: In this example, we take “矿井提升机” (mine hoist) as the input data. The constructed Trie tree contains the following five feature words: “<矿井> (mine)”, “<提升机> (mine lifting)”, “<提升> (lifting)”, “<提机> (hoist)”, and “<矿井提升机> (mine hoist)”.

Individual characters, like “矿”, “井”, “提”, “升”, and “机”, will be matched with feature words in the Trie tree using a dictionary adapter. Such a matching technique allows for a precise association between characters and their relevant feature words, enhancing the efficiency and accuracy of entity recognition in a specific domain. L represents the depth of the coding layer and k refers to the k-th coding layer.
With these three carefully designed components, the DdERT model can efficiently process and parse textual information in the mine hoist fault domain, thus achieving accurate named entity recognition.

3.1. Algorithm for Building Domain Dictionary

To enhance the efficiency and accuracy of the DdERT model in entity recognition, we developed a domain-specific dictionary tailored for the mine hoist equipment sector. This dictionary was constructed based on various mining equipment documents, including operation manuals, safety standards, and malfunction records, extracting relevant feature words.

For efficient storage and retrieval of these feature words, we utilized the Trie data structure. The strength of a Trie lies in its ability to share common prefixes, substantially reducing both storage requirements and retrieval time. Each node in the Trie represents a character, with the path from the root to a specific node forming a string.

Integrating this domain-specific dictionary into the DdERT model amplifies the model’s text-processing capabilities and enables it to gain in-depth expertise about mine hoist equipment. This approach significantly elevates the model’s performance in specialized domains. To solve this issue, we designed Algorithm 1.

**Algorithm 1:** Dictionary tree construction algorithm

**Input:** $s_c = \{c_1, c_2, \ldots, c_n\}$, a sentence with $n$ characters $s_c$

**Output:** $s_{cw} = \{(c_1, w_{s1}), (c_2, w_{s2}), \ldots, (c_n, w_{sn})\}$

1. Initialize the word list $W_{s} = \{w_{s1}, w_{s2}, \ldots, w_{sn}\}$
2. Initialize the dictionary tree root node
3. for $i \leftarrow 1$ to $n$ do
4. repeat
5. Iterate through the word list $W_{s}$ to obtain the words matching the matching words $w_{s_i}$
6. Take $(c_i, w_{s_i})$ constructing dictionary tree sub-nodes
7. until the traversal is complete $W_{s}$ stop
8. end for

3.2. The Dictionary-Enhancement BERT

In processing text data related to mine hoists, generating word vectors for feature words within the DdERT model is pivotal. Here are the key steps and processes:

1. Subsequence Matching: For a given input sentence $S$, all character subsequences are traversed and matched against a pre-constructed dictionary tree (feature word repository) using a dictionary fusion algorithm. Each character can align with up to three feature words, with any shortfall filled by the special symbol <PAD>.

2. BERT Embedding: The dictionary-augmented BERT model embeds the character-word pairs, generating the corresponding word vectors.

3. Transformer Processing: The word vectors from the previous step are fed into BERT’s transformer encoders. To incorporate domain-specific insights, a dictionary adapter is interposed between consecutive layers of the transformer, aiming to infuse domain-specific lexical information.

4. Lexicon Adaptation: A dictionary adapter performs transformations upon specific layers’ outputs, integrating lexical insights into the feature vector.

5. Subsequent Processing: Post-transformation, these feature vectors proceed through additional transformer layers, allowing the model to comprehend better and represent the textual data.

This methodology synergizes BERT’s profound text processing capability with domain-specific dictionary insights, optimizing entity recognition accuracy and efficiency within niche domains, such as mine hoisting equipment. The process is shown in Figure 1, Character Encoding Layer.
3.3. Dictionary Fusion Algorithm

In the DdERT architecture, dictionary adapters are introduced to enhance the model’s performance on entity recognition tasks. The dictionary adapter is nested between two transformer encoder layers and aims to fuse word-level and word-level feature representations finely.

Character-level and word-level feature representations have different granularities and have potential advantages in entity boundary detection and recognition. Character-level features capture fine-grained information, while word-level features provide richer semantic information. The dictionary adapter balances these two representations by dynamic weight assignment, prioritizing word-level features that match better with word-level features. In order to efficiently retrieve and match the domain feature words, the pre-built dictionary tree is used to calculate the similarity between the input word and the domain feature words. To solve this issue, we designed Algorithm 2.

Algorithm 2: Dictionary Fusion algorithm

Input: \( WS = \{w_{s1}, w_{s2}, \ldots, w_{sn}\} \), \( Y = \{y_{c1}, y_{c2}, \ldots, y_{cn}\} \)

Output: \( H = \{h_1, h_2, \ldots, h_n\} \)

1. for \( i \leftarrow 1 \) to \( n \) do
   for \( j \leftarrow 1 \) to \( n \) do
     \( w_{ij} = w_{sij} \)
     \( x^{ws}_{ij} = e^{w_{ij}} \)
   end for
   end for

2. A nonlinear transformation \( v^{ws}_{ij} = W_2(\text{tanh}(W_1x^{ws}_{ij} + b_1)) + b_2 \) of the word embedding vector \( x^{ws}_{ij} \), yields \( V_i = \{v^{ws}_{i1}, v^{ws}_{i2}, \ldots, v^{ws}_{in}\} \)

3. for \( i \leftarrow 1 \) to \( n \) do

4. Calculating word and domain feature word correlation \( a_i = \text{softmax}(y_i^T W_{\text{attn}} V_i^T) \)

5. Calculate the weighted sum of all words \( z^w_i = \sum_{j=1}^{n} a_{ij}v^{ws}_{ij} \)

6. Weighted lexical information injection character vector \( \hat{h}_i = y^T_i + z^w_i \)

7. end for

3.4. Decoding Layer

After designing the input layer and encoding layer, to capture the dependencies between consecutive labels in the sequence labeling, a CRF was introduced into the model as the decoding layer.

(1) Calculating probability scores

For the hidden output of the last layer \( H^L = \{h^L_1, h^L_2, \ldots, h^L_n\} \), a probability score \( p \) is calculated for each label. This score will be used as an input to the CRF layer to estimate the probability of the label sequence.

\[
p(y|s) = \frac{\exp\left(\sum_i \left( O_{iy_i} + T_{y_{i-1}y_i} \right) \right)}{\sum_y \exp\left(\sum_i \left( O_{iy_i} + T_{y_{i-1}y_i} \right) \right)} \tag{1}
\]

Among them, \( T \) is the transition score matrix, which represents the transition probability between labels. \( y \) denotes the sequence of labels at all positions.

(2) The probability of a sequence of labels

The probability of a label sequence \( y = \{y_1, y_2, \ldots, y_n\} \) is defined as follows:

\[
O = W_y H^L + b_y \tag{2}
\]

(3) Loss function
Given \( N \) labeled data \( \{s_j, y_j\}_{j=1}^N \), we trained the model by minimizing the negative log-likelihood loss at the sentence level. This objective function can be expressed as follows.

\[
L = -\sum_j \log(p(y_j|s_j))
\]

In decoding, we apply the Viterbi algorithm to find the tag sequence with the highest score. The Viterbi algorithm is a dynamic programming algorithm that optimizes the score of the whole sequence by finding the most likely partial sequence at each time step.

In summary, as a decoding layer, the CRF layer can effectively simulate the label sequence’s dependencies and, combined with the Viterbi decoding algorithm, enables the model to show higher accuracy on the sequence labeling task. By combining CRF with deep neural networks, our model can learn more complex and refined feature representations and complex relationships between labels.

4. Dataset
4.1. Data Analysis
4.1.1. Design of Fault Diagnosis Ontology Model Structure

As a formal structure describing entities, attributes, and relationships, ontology has been introduced into various intelligent fields to realize information fusion and sharing in recent years. “Ontology is the explicit formal specification of the shared conceptual model”, which points out that ontology has four levels of meaning: conceptual model, explicit, formal, and shared. Ontology first appeared in the field of philosophy.

The skeleton method \([32,33]\), TOVE method \([34]\), methodology method, and seven-step method \([35]\) are widely used and representative ontology modeling methods. After application by many scholars, combined with the advantages and disadvantages of the above ontology modeling methods, this work uses the seven-step method to build the mine hoist fault domain ontology model, and the specific process is shown in Figure 2.

![Figure 2. Structure of Fault Diagnosis Ontology Model.](image)

4.1.2. Data Entity Type

Based on the current national standards GB/T 35737-2017 \([36]\) and GB/T 20961-2018 \([37]\), the entity types of mine hoists were divided into fault phenomenon, fault cause, fault location, and fault effect and repair measures (see Table 1 for a description of the specific types).
Table 1. Entity types.

<table>
<thead>
<tr>
<th>Entity Type</th>
<th>Illustrate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fault Phenomenon</td>
<td>External manifestations of faults</td>
</tr>
<tr>
<td>Repair Measures</td>
<td>Various measures taken to restore the fault to its original usable state</td>
</tr>
<tr>
<td>Fault Effect</td>
<td>Result of a failure mode on the use, function, or state of a product</td>
</tr>
<tr>
<td>Fault Cause</td>
<td>Factors related to design, manufacturing, use, and maintenance that cause malfunctions</td>
</tr>
<tr>
<td>Fault Location</td>
<td>The location where the fault occurred</td>
</tr>
</tbody>
</table>

4.2. Data Sources

This study aims at entity recognition in the field of mine hoist fault. As there is no public data set in this field, fault text information was extracted from the on-site maintenance log, operation manual, safety regulations, spot inspection table, and other information of a state-owned large non-ferrous metal group, and literature related to mine hoists was collected from a certain country’s National Knowledge Infrastructure, Wanfang Data, Baidu Baike, and other websites. By preprocessing the collected data, such as removing stop words, deleting unrecognizable special symbols and useless spaces between strings, etc., a 7733 text corpus in the field of mine hoist was obtained. Secondly, the LabelStudio annotation platform was used for semi-automatic annotation and manual proofreading of the corpus. The corpus contained a total of 7733 entities, including 2282 fault phenomenon entities, 3749 fault location entities, 791 fault cause entities, 110 fault effect entities, and 801 repair measures entities. The statistics of the distribution and the percentage of entities are shown in Figure 3.

![Entity Distribution Chart](image)

Figure 3. Entity Distribution Chart.
This dataset not only pioneered research in the field of mine hoist faults, but also reflected the complexity of the context in which relationships occur in the text. Thus, the dataset is annotated by B-I-O (Begin–Inside–Outside) method, such as the sentence “improper gear meshing leads to high oil temperature”, where “gear” is “fault location”, which is marked as B-FLN. “Improper meshing” is the “fault cause”, marked as B-FCE; “oil temperature is too high” is the “fault phenomenon”, marked as B-FPN. Information not related to the entity is labeled as O. An example of the annotation is shown in Figure 4. Finally, the dataset for mine hoist fault named entity recognition was constructed and divided into 6182 training sets, 775 validation sets, and 775 test sets according to 80%, 10%, and 10%, respectively.

5. Experiment

5.1. Evaluating Indicator

In this study, the model was constructed utilizing the PyTorch framework, leveraging the Chinese pre-trained language model BERT as the underlying encoder, paired with the Adam optimizer for optimization. It is pertinent to note that a consistent set of training parameters was employed across all models to ensure a fair and balanced comparison during the experimental evaluations. These parameters, delineated in Tables 2 and 3, were determined through meticulous parameter tuning experiments conducted on the validation set.

<table>
<thead>
<tr>
<th>Table 2. Model preset parameters.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Parameter Name</strong></td>
</tr>
<tr>
<td>Word2vec word vector dimensionality</td>
</tr>
<tr>
<td>Learning rate</td>
</tr>
<tr>
<td>Batch size</td>
</tr>
<tr>
<td>Random discard rate</td>
</tr>
<tr>
<td>Hidden layer dimension</td>
</tr>
<tr>
<td>Number of attention mechanism heads</td>
</tr>
<tr>
<td>Maximum sentence length</td>
</tr>
<tr>
<td>Maximum fusion of vocabulary information per Chinese character</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 3. Experimental environment parameters.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Environment</strong></td>
</tr>
<tr>
<td>GPU</td>
</tr>
<tr>
<td>CPU</td>
</tr>
<tr>
<td>Memory</td>
</tr>
<tr>
<td>Programming language</td>
</tr>
<tr>
<td>Training framework</td>
</tr>
</tbody>
</table>

5.2. Experimental Parameters

Precision (P), recall (R), and F1-score were used as evaluation indexes in this experiment. Precision is the percentage of system results correctly identified and recall is the percentage of total entities correctly identified by the system. The F1-score is the weighted
The harmonic mean of precision and recall and evaluates the overall performance of the model. The calculation methods of each evaluation index are shown in Formulas (4)–(6).

$$\text{Precision} = \frac{TP}{TP + FP}$$ (4)

$$\text{Recall} = \frac{TP}{TP + FN}$$ (5)

$$F1 - \text{score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$ (6)

where TP represents the number of sample data that the test set model can correctly predict, FP represents the number of irrelevant sample data that the test set model predicts, and FN represents the number of sample data that the test set model predicts unsuccessfully.

5.3. Comparative Experimental

In order to verify the entity recognition effect of the DdERT model, we conducted comparative experiments using a consistent set of training parameters across all models, ensuring a fair and balanced evaluation. The details of these parameters are illustrated in Tables 2 and 3. Five types of entity are identified on the self-built mine hoist data set, and the BiLSTM model, BiLSTM-CRF model, BERT model, BERT-BiLSTM-CRF model, and LEBERT model are selected for comparative experiments.

As shown in Figure 5, the entity recognition effect of the DdERT model on the mine hoist dataset was significantly better than those of the BiLSTM, BiLSTM-CRF, BERT, and BERT-BiLSTM-CRF models. Due to the particularity of the description of the fault of the hoist, it was easy to cause different types of entities due to different contexts. For example, “reducer” in “hoist reducer” means equipment components, and “reducer must be reliably grounded, avoid welding work on site, prevent gear and bearing discharge, damage gear and bearing” means repair measures. Different entity types make traditional static word vectors perform poorly. As the number of entities and the complexity of entities in the fault description text were higher than those in the general text, the defects of the BiLSTM method lacking local feature perception and information loss were more evident in the elevator data set. However, by introducing BERT, compared with the BiLSTM and BiLSTM-CRF models, the recognition accuracy, recall rate, and F1 value were improved, effectively solving the problem of entity polysemy. After adding word features to text embedding, the LEBERT model improved the recognition accuracy by 7.78% and the F1 value by 3.37% compared with BERT for the hoist dataset. The DdERT model was more sensitive to the word-level semantic changes caused by domain professional terms by constructing a domain dictionary. Compared with LEBERT, the recognition accuracy was increased by 5.83% and the F1 value was increased by 7.09%. It can be seen that the entity recognition effect of the DdERT model was more evident when integrating the domain dictionary and using the CRF algorithm to extract the global optimal labeling sequence, so the DdERT model was more suitable for the entity recognition task in the mine hoist domain.
ensuring a fair and balanced evaluation. The details of these parameters are illustrated in Tables 2 and 3. Five types of entity are identified on the self-built mine hoist data set, and the BiLSTM model, BiLSTM-CRF model, BERT model, BERT-BiLSTM-CRF model, and LEBERT model are selected for comparative experiments.

As shown in Figure 5, the entity recognition effect of the DdERT model on the mine hoist dataset was significantly better than those of the BiLSTM, BiLSTM-CRF, BERT, and BERT-BiLSTM-CRF models. Due to the particularity of the description of the fault of the hoist, it was easy to cause different types of entities due to different contexts. For example, "reducer" in "hoist reducer" means equipment components, and "reducer must be reliably grounded, avoid welding work on site, prevent gear and bearing discharge, damage gear and bearing" means repair measures. Different entity types make traditional static word vectors perform poorly. As the number of entities and the complexity of entities in the fault description text were higher than those in the general text, the defects of the BiLSTM method lacking local feature perception and information loss were more evident in the elevator data set. However, by introducing BERT, compared with the BiLSTM and BiLSTM-CRF models, the recognition accuracy, recall rate, and F1 value were improved, effectively solving the problem of entity polysemy. After adding word features to text embedding, the LEBERT model improved the recognition accuracy by 7.78% and the F1 value by 3.37% compared with BERT for the hoist dataset. The DdERT model was more sensitive to the word-level semantic changes caused by domain professional terms by constructing a domain dictionary. Compared with LEBERT, the recognition accuracy was increased by 5.83% and the F1 value was increased by 7.09%. It can be seen that the entity recognition effect of the DdERT model was more evident when integrating the domain dictionary and using the CRF algorithm to extract the global optimal labeling sequence, so the DdERT model was more suitable for the entity recognition task in the mine hoist domain.

5.4. Comparative Experimental

In order to verify the recognition performance of the DdERT model for various entity types of mine hoists in the same experimental environment, the DdERT model achieved the optimal F1 value on the data set. Figure 6 shows the recognition results of the DdERT model on the five types of entities of the mine hoist data set in this work.

As shown in the table, the fault location recognition effect of the DdERT model was significantly better than that of other categories. As the number of fault location entity labels was the largest and the entity features were apparent, it had a good recognition effect, and its F1 value reached 98.91%. However, the recognition effect of fault-affected entities was the worst, with an F1 value of only 88.37%. This was mainly due to the small number of samples of this kind of entity with variable types and rich semantics, and the model could not thoroughly learn the characteristics of this kind of entity, resulting in a poor recognition effect.

Figure 5. Comparison model results.

Figure 6. Comparison category results of the model.
As shown in the table, the fault location recognition effect of the DdERT model was significantly better than that of other categories. As the number of fault location entity labels was the largest and the entity features were apparent, it had a good recognition effect, and its F1 value reached 98.91%. However, the recognition effect of fault-affected entities was the worst, with an F1 value of only 88.37%. This was mainly due to the small number of samples of this kind of entity with variable types and rich semantics, and the model could not thoroughly learn the characteristics of this kind of entity, resulting in a poor recognition effect.

6. Conclusions

Information extraction is an indispensable part of knowledge graph construction. This study proposes a DdERT model for entity recognition tasks in the mine hoist fault domain based on the BERT pre-trained language model. Aiming at the problem that the recognition effect of rare entities and equipment model abbreviations in mine hoists data is not good, lexical information is introduced to provide more semantic information and entity boundary information, and a vocabulary fusion method based on a domain dictionary is designed so that the model can process multiple sentences at the same time. The experiments show that the model in this study had good recognition results in the mine hoist fault data set and had higher precision and recall rate. This will support the construction of the mine hoist’s fault domain knowledge graph. In the next step, this model will continue to be optimized, and the fault domain knowledge graph’s construction will be gradually realized.

Author Contributions: Proposal of research topics, X.D. (Xiaochao Dang) and L.W.; Data processing and experiments, L.W. and H.D.; Writing—original draft preparation, L.W. and H.D.; Writing—review and editing, X.D. (Xiaohui Dong), L.W. and F.L.; supervision, X.D. (Xiaochao Dang); project administration, X.D. (Xiaochao Dang); funding acquisition, X.D. (Xiaochao Dang) All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Natural Science Foundation of China (Grant 62162056), Industrial Support Foundations of Gansu (Grant No. 2021CYZC-06) by X.D. (Xiaochao Dang).

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

7. Lv, K.; Gao, C.; Si, J.; Feng, H.; Cao, W. Fault Coil Location of Inter-Turn Short-Circuit for Direct-Drive Permanent Magnet Synchronous Motor Using Knowledge Graph. IET Electr. Power Appl. 2020, 14, 1712–1721. [CrossRef]


Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.