DOI: 10.19884/j.1672-5220.202403019

Exploring Techniques for Building Language Models Targeted at Sewing Equipment Operation and Maintenance Management

LIU Bing[1], LIU Ying[1], ZHENG Xiaohu[2*], LI Xiechen[3], DU Siqi[4]

1. Hangzhou Zhongfu Technology & Innovation Co.,Ltd., Hangzhou 311103, China
2. Institute of Artificial Intelligence, Donghua University, Shanghai 201620, China
3. College of Information Science and Technology, Donghua University, Shanghai 201620, China
4. College of Mechanical Engineering, Donghua University, Shanghai 201620, China

Abstract: The intelligent operation and maintenance management of sewing equipment needs to solve the problem of information mining and language model construction of unstructured text, which is of great significance to improve the speed and accuracy of the diagnosis of equipment defects and faults, and realize the intelligent decision-making of equipment maintenance. In this paper, firstly, we propose a method based on bidirectional encoder representations from transformers-conditional random fields (BERT-CRF) to extract key entity information, such as device names and attributes. Then, through the relationship extraction model based on bidirectional gated recurrent unit-attention (BiGRU-Attention), the semantic association between entities is captured effectively to provide support for the construction of the sewing equipment knowledge graph (KG). According to the text analysis scenario of sewing equipment, the model is specially trained and optimized in the task scenarios of text entity recognition, information extraction and fault diagnosis of sewing equipment. Compared with existing deep learning algorithms, the proposed method achieves a 20% to 30% performance improvement on the validation and test sets, demonstrating significant advantages in the recall rate and the accuracy. To facilitate the mining of unstructured text information on sewing equipment, this study provides a reference for constructing a KG that integrates data on flat sewing equipment, including aspects of equipment fault operation, maintenance and flat sewing process route design.

Key words: sewing equipment operation management; language model; entity extraction; relation extraction; knowledge graph

CLC number: TP391.1; TS103 **Document code**: A

Article ID: 1672-5220(2024)03-0315-08

Open Science Identity (OSID)

0 Introduction

In modern manufacturing, sewing equipment serves as a critical tool for achieving automation in garment production, and the optimization of its performance and technological innovation are vital for enhancing production efficiency and product quality. With the widespread application of artificial intelligence and big data technologies, there is an increasing demand for intelligent management and maintenance of sewing equipment[1]. Non-structured data from sewing equipment provides crucial reference information for monitoring equipment condition, analyzing defects, diagnosing failures and managing the entire lifecycle of the equipment[2]. Therefore, effectively mining non-structured text information plays a significant role in improving the speed and accuracy of diagnosing equipment defects and failures, thereby assisting in maintenance decision-making[3-4].

Sewing equipment knowledge possesses distinct characteristics such as specialization, complexity and non-structured nature. Traditional equipment management has heavily relied on human expertise and skills for operation and maintenance. This approach is often inefficient and unsuitable for real-time monitoring. Knowledge graphs (KGs), however, can represent knowledge within the sewing equipment domain in a structured graph format, converting unstructured textual materials into searchable and inferable knowledge bases, enabling efficient and accurate operation and maintenance management of sewing equipment[5].

In the realm of intelligent applications of KGs, graph databases, exemplified by Neo4j, serve as the principal repositories for knowledge storage. These databases enable core scenarios such as semantic retrieval, semantic question answering (Q&A), knowledge visualization and analysis, as well as

Received date: 2024-03-29

* Correspondence should be addressed to ZHENG Xiaohu, email: xhzheng@dhu.edu.cn

Citation: LIU B, LIU Y, ZHENG X H, et al. Exploring techniques for building language models targeted at sewing equipment operation and maintenance management [J]. *Journal of Donghua University (English Edition)*, 2024, 41(3): 315-322.

knowledge recommendation[6]. Concurrently, Hedberg et al.[7] harnessed KG technology in manufacturing to furnish comprehensive product lifecycle traceability, while Huet et al.[8] devised a design rule recommendation system grounded in KG, and Hao et al.[9] proposed a mechanical component design tool facilitating the generation of design solutions. However, the application of KG technology in the sewing industry remains scarce, rendering the holistic integration of information generated across various stages of the garment sewing process challenging.

In order to effectively manage the large amount of information generated during the sewing process, this study proposes a KG construction method of sewing process information fusion. Language models serve as a central component in constructing KG, capable of generating reasonable text based on given contexts, thus contributing to the creation of richer and more accurate KG. By training on large-scale corpora, language models can learn about device entities, attributes and relationships among them, subsequently generating knowledge bases that reflect actual scenarios[10]. We use natural language processing (NLP) technology, especially the entity extraction model based on bidirectional encoder representations from transformers-conditional random fields (BERT-CRF) and relationship extraction model based on bidirectional gated recurrent unit-attention (BiGRU-Attention), to dive deeply into and leverage the non-structured text related to sewing equipment. These models accurately extract key entities and their attributes from the text and comprehend the complex relationships between entities, laying a solid foundation for building a high-quality KG for sewing equipment. Such efforts not only enhance production efficiency and product quality but also significantly reduce maintenance costs, providing effective technical support for the digital transformation of manufacturing industries.

1 Entity Extraction Model Based on BERT-CRF

Entity extraction involves using NLP technology and machine learning algorithms to identify and extract meaningful units with specific meanings from texts, such as names of sewing equipment, types of malfunctions and methods for troubleshooting[11]. The process annotates these entities, as demonstrated in Fig. 1. From the input text, the result of entity extraction is that the sewing equipment is "HX68S series", its equipment category is "overlock sewing machine", the fault description is "suture breakage", the fault code is "E11", the fault handing method is "replace the sewing thread". The result of the relationship extraction is "E11", which means that the fault is "suture breakage" and the corresponding fault handing method is "replace the sewing thread". This endeavor often relies on a variety of entity extraction techniques, including hidden Markov models (HMM), conditional random fields (CRF) and bidirectional long short-term memory (BiLSTM) networks, to enhance both accuracy and efficiency in the extraction process[12]. These technologies have been widely applied in areas like translation and Q & A. However, they still tend to be limited in their capacity to analyze complex relationships between words and phrases, demonstrating less than optimal precision in understanding the deeper semantics within the text.

The entity extraction method proposed in this paper utilizes a sequence labeling model combing a pre-trained language model bidirectional encoder representations from transformers (BERT) with a CRF. This model employs BERT as the encoder to extract feature representations from input sequences which are then fed into a CRF layer for labeling. The CRF layer leverages contextual information to jointly model labels, thereby enhancing the accuracy of the tagging results.

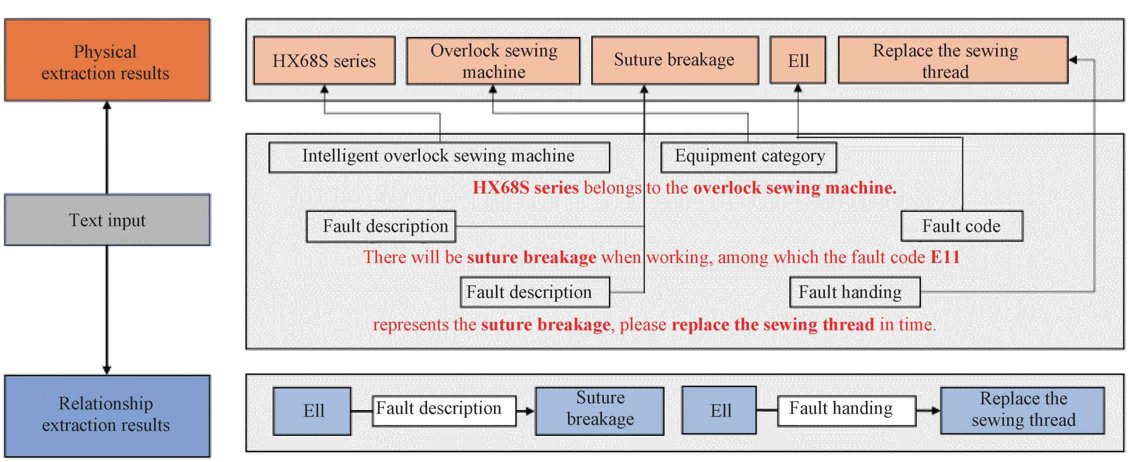

Fig. 1 General process of KG construction

1.1 BERT model

BERT[13-15] is a bidirectional pre-trained language model based on the transformer architecture, which learns rich semantic and contextual information by pre-training on a large corpus of unlabeled text data. The BERT model encodes each word in the input text into a contextually relevant vector representation which can be used as input features for various NLP tasks.

The input to BERT is the direct sum of three types of embeddings, as illustrated in Fig. 2. Token embeddings are word or character vectors; position embeddings are positional vectors that capture the position of each token in the sequence; segmentation embeddings differentiate between different segments in the input, such as when dealing with multiple sentences or paragraphs.

A special token "[CLS]" is used as the starting marker for a sequence, while the token "[SEP]" is employed to segment paragraphs or separate distinct segments within the input.

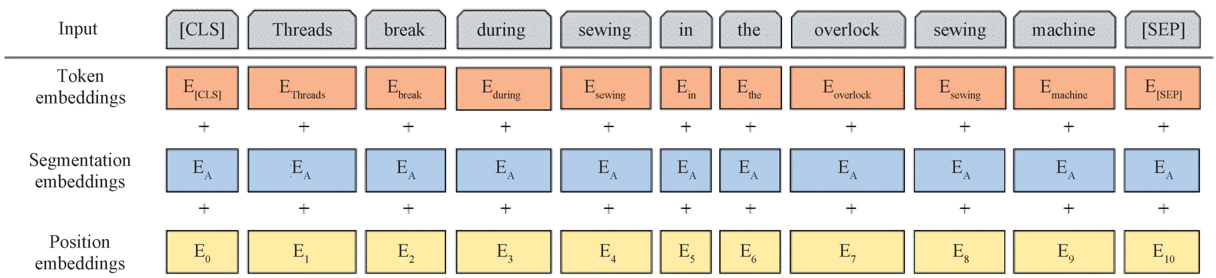

Fig. 2 BERT input

BERT adopts a multi-layer bidirectional transformer encoder that fine-tunes on tasks, allowing each character within a sentence to encode information from any other character, integrating context from both left and right sides. Each block within the encoder comprises a multi-head self-attention mechanism and a fully connected feed-forward network. The multi-head self-attention means performing multiple attention computations, where each attention head focuses on different aspects of the information within the sentence. The outputs from all attention heads are then concatenated together

$$MultiHead(\boldsymbol{Q}, \boldsymbol{K}, \boldsymbol{V}) = Concat(h_1, h_2, \cdots, h_h)\boldsymbol{W}_0, \quad (1)$$

$$h_i = Attention(\boldsymbol{QW}_{i,Q}, \boldsymbol{KW}_{i,K}, \boldsymbol{VW}_{i,V}), \quad (2)$$

where $\boldsymbol{Q}$ is the query matrix; $\boldsymbol{K}$ is the key matrix; $\boldsymbol{V}$ is the value matrix; W_Q, W_K and W_V are the weight matrices; h stands for the head in the multihead self-attention; i means the number of self-attention used.

Self-attention operates through computations involving three matrices $\boldsymbol{Q}$, $\boldsymbol{K}$ and $\boldsymbol{V}$, where it takes each word vector input into the encoder and performs dot products followed by weighted summation across the entire input sequence to derive the output result at that particular position, as expressed:

$$Attention(\boldsymbol{Q}, \boldsymbol{K}, \boldsymbol{V}) = softmax\left(\frac{\boldsymbol{QK}^{\mathrm{T}}}{\sqrt{d_k}}\right)\boldsymbol{V}. \quad (3)$$

Each block in the decoder, unlike the encoder, includes an additional encoder-decoder attention mechanism, which takes as inputs both the decoder's own input and the encoder's output.

Due to the inherent property of self-attention which only focuses on each word itself, in order to train a bidirectional transformer model, a certain proportion of words must be randomly masked out during pre-training. These masked words are either replaced with a MASK token or another word, while the remaining words remain unchanged. The model is then trained to predict the correct masked words based on the surrounding context, thereby integrating relevant information from the context.

1.2 CRF model

For sequence labeling problems, a softmax classifier is typically used during the prediction phase to classify tags. However, the softmax classifier does not take into account the dependencies between labels, which are inherent in sequence labeling tasks[16]. On the contrary, a CRF can utilize a log-linear model to represent the joint probability of an entire feature sequence, making it better suited for predicting labels in sequence labeling.

Assuming the length of a sentence is n, and the sentence sequence is $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$, with the corresponding predicted label sequence being $\boldsymbol{y} = (y_1, y_2, \cdots, y_n)$, the total score for the predicted sequence would be calculated:

$$s(\boldsymbol{x}, \boldsymbol{y}) = \sum_{i=0}^{n} \lambda_{y_i, y_{i+1}} + \sum_{i=1}^{n} P_{i, y_i}, \quad (4)$$

where s represents score; λ represents the transfer score between labels; P_{i, y_i} represents the score of each word to the corresponding y_i label.

Since the prediction sequence has many possibilities, only one of which is the most correct, all possible sequences should be globally normalized to produce the probability from the original sequence to the predicted sequence:

$$\boldsymbol{P}(y \mid x) = \frac{\mathrm{e}^{s(x, y)}}{\sum_{\tilde{y} \in y_x} \mathrm{e}^{s(x, \tilde{y})}}, \quad (5)$$

where $\boldsymbol{P}$ is the score matrix output from the attention layer; $\bar{y}$ denotes the real labeling sequence; y_x means all possible dimension sequences.

1.3 BERT-CRF model

In traditional neural network models for NLP, the use of NLP tools to preprocess text corpora, such as tokenization, where the segmented words serve as input vectors, can lead to the accumulation and propagation of errors, resulting in inaccurate predictions of labels. To mitigate this error propagation issue, BIO sequence labeling rules are employed, where "B" denotes the beginning of an event trigger word, "I" signifies the inside of an event trigger word, and "O" indicates a non-event trigger word. The model's inputs include token embeddings, position embeddings and segmentation embeddings, with the output being the labeling result for each character.

The BERT-CRF model is an extension of the BERT model that incorporates a CRF linear layer at its end, as depicted in Fig. 3. Trm in Fig. 3 represents the transformer. This configuration allows the model to jointly predict the label sequence considering the dependencies between adjacent tokens, thereby enhancing the coherence and accuracy of the sequence labeling task.

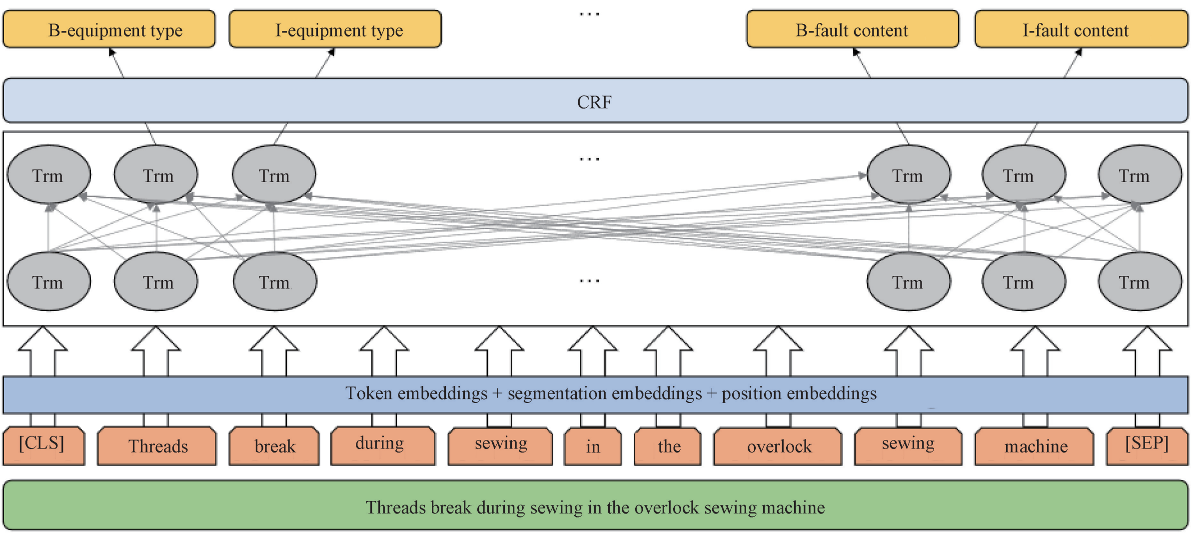

Fig. 3 Structure diagram of BERT-CRF model

2 Relationship Extraction Model Based on BiGRU-Attention

Relationship extraction builds upon entity recognition by further identifying and extracting specific connections between entities, such as the association between devices and their malfunctions or the interactions among components of a device. Its functionality is illustrated in Fig. 1. This step is pivotal for constructing a rich and multi-dimensional KG. A model combining convolutional neural networks (CNNs) and attention mechanisms is adopted to automatically identify various types of relationships between entities from text data. CNNs excel in capturing local features in textual data, while the attention mechanism automatically focuses on the information most relevant to relationship extraction, thus enhancing both accuracy and efficiency in relationship identification[17].

As a critical step in building a KG, relationship extraction significantly impacts the quality and usability of the constructed graph. Traditional methods relying on pattern matching for relationship extraction incur high computational costs and are prone to error propagation, leading to poor generalizability. In response to these challenges, this paper proposes the BiGRU-Attention model for relationship extraction, as shown in Fig. 4. The approach boasts fast convergence speed and improves the precision of relationship extraction.

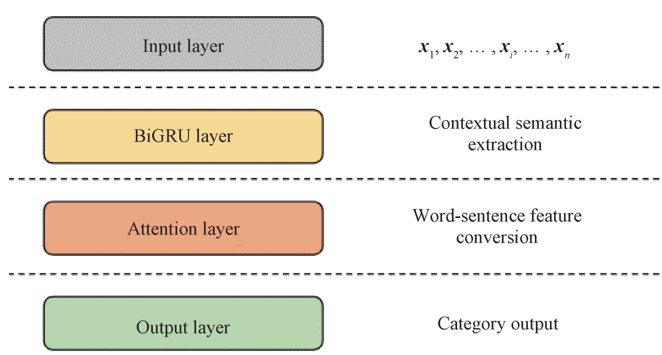

Fig. 4 Structural diagram of BiGRU-Attention model

BiGRU[18], standing for bidirectional gated recurrent unit, represents a recurrent neural network structure capable of capturing contextual information within sequential data. By concurrently operating through forward and backward GRU units, BiGRU discerns a more comprehensive range of semantic information within sequential data. Attention, on the other hand, refers to assigning weights to detected sequences in the input text information, enabling the model to focus more intently on features most pertinent to relationship extraction.

In the BiGRU-Attention model, the processed self-built sewing technology knowledge corpus is embedded into a vector form, with the resulting word vectors serving as input to the BiGRU. Through the BiGRU process, contextual semantic features are captured. Simultaneously, the integration of an attention layer facilitates the extraction of relevant features. Ultimately, the softmax function is utilized for classification, yielding the optimal results.

1) Input layer. The input data consist of data from a self-built sewing technology knowledge corpus, which is first transformed into word embeddings. This means that each character or word is represented as a fixed-dimensional vector, allowing the model to better understand and process text data.

2) BiGRU layer. This layer performs forward and backward encoding on the fault statements, enabling the model to capture contextual information in the text. This enhances the understanding of the meaning of each character or word within a sentence.

3) Attention layer. It assigns certain weights to the sequences detected in the input text. This attention mechanism helps in further extracting features from the fault sentences, mitigating the negative impact of noise data, and thereby improving the model's performance in relation extraction.

4) Output layer. Through the combined action of the BiGRU neural network and the softmax activation function, the output layer delivers the relationship type.

3 Experimental Verification

3.1 Entity and attribute extraction results

The entity extraction task primarily involves content related to maintenance and operation issues of sewing equipment. According to the constructed ontology model, the entities to be extracted encompass categories of equipment, specific equipment models, details of malfunctions, troubleshooting procedures, malfunction codes, and causes of failure. The visualization of the ontology model underlying the entity extraction process is depicted in Fig. 5.

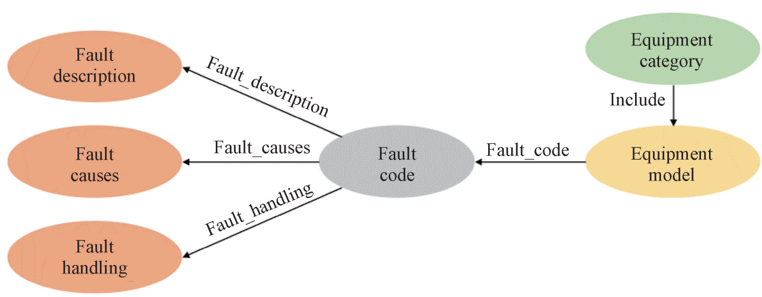

Fig. 5 Recommended knowledge system ontology model

Taking "fault code" and "fault description" as examples, in the context of the entity extraction task, the "fault code" would be considered the "head entity", while the "fault description" serves as the "tail entity". In the subsequent relationship extraction task, the combination of these two, expressed as "fault_description", constitutes the "relation entity".

From sewing technique documents and sewing equipment operation manuals, we have partitioned out 658 data entries, totaling 9 878 characters. Based on the ontology constructed earlier, using the previously described BIO three-character tagging method, each character in the dataset has been annotated. A total of 423 entities have been tagged, and the results of the entity annotations are presented in Table 1.

Table 1 Entity annotation

Connotation	Annotation	Quantity
Equipment category	B-equipment category, I-equipment category	5
Equipment model	B-equipment model, I-equipment model	50
Fault description	B-fault description, I-fault description	58
Fault handling	B-fault handling, I-fault handling	72
Fault causes	B-fault causes, I-fault causes	56
Fault code	B-fault code, I-fault code	182

After annotating the documents to form a dataset, they are divided into training, test and validation sets at a ratio of 8 : 1 : 1. Using the training set, cache files for the vocabulary and label tables are generated. The text is then equally segmented, with every 50 characters forming a group. Subsequently, preprocessing is performed where the top 5 000 most frequently occurring characters are retained, and all other less frequent words, numerals, and alignment-padding terms are respectively replaced by <UNK>, <NUM> and <PAD>.

A BERT-CRF model is trained on the preprocessed annotated training set for entity extraction. The parameters of the entity extraction model are detailed in Table 2.

Table 2 Entity extraction model parameters

Parameter	Value
Continuous bag of words model window size	2
Number of heads of attention mechanism	12
Word embedding vector dimension	300
Learning rate	1×10^{-3}
Batch size	64
Dropout	0.3
Iterative algebra	100

In order to evaluate the performance of the BERT-CRF model in recognizing named entities within the context of sewing equipment fault maintenance corpus, comparative experiments were conducted using RNN-CRF, LSTM-CRF and BiLSTM-CRF models, as well as the BERT-CRF model proposed in this study. The experimental results are presented in Table 3.

Table 3 Comparison of experimental results of different models

Model	Accuracy rate/%	Recall rate/%	F1 score/%
RNN-CRF	82.72	80.90	83.80
LSTM-CRF	85.65	83.78	86.19
BiLSTM-CRF	88.45	86.52	88.47
BERT-CRF	91.82	90.44	91.01

Recall rate refers to the proportion of samples that are actually positive examples predicted by the model to be positive examples. The F1 score is the harmonic mean of accuracy and recall, and its value ranges from 0 to 1. The higher the F1 score, the better the performance of the model and the better the balance between accuracy and recall.

As indicated by Table 3, the RNN-CRF model suffers from the vanishing gradient problem which hinders its ability to handle long-term dependencies, leading to a relatively lower accuracy rate of 82.72% for the named entity recognition task. The LSTM-CRF model, compared to the RNN-CRF model, improves upon this issue by incorporating input gates, forget gates and output gates, thereby capturing relations across the entire sentence and achieving an accuracy of 85.65%. The BiLSTM-CRF model, utilizing bidirectional LSTMs, further boosts the accuracy to 88.45%.

The BERT-CRF model surpasses these three models due to its parallel capability of capturing global context information. In comparison, it achieves the best performance, with the highest accuracy rate, recall rate and F1 score of 91.82%, 90.44% and 91.01%, respectively. When compared with the previous most effective model (BiLSTM-CRF), the BERT-CRF model registers enhancements of 3.37 percentage points in the accuracy rate, 3.92 percentage points in the recall rate, and 2.54 percentage points in the F1 score.

3.2 Relation extraction result

The relationship extraction task builds upon the foundation of entity extraction, proceeding with relational annotation. The division of the training and test sets for this task follows the same criteria as those used in the entity extraction task. As per the fault maintenance knowledge ontology model illustrated in Fig. 5, key relationships to be extracted include "fault_description", "fault_handling" and "fault_cause". Part of the training set corpus used to develop the relationship extraction model for sewing equipment fault maintenance is presented in Table 4.

Table 4 Corpus of relationship extraction model

No.	Head entity	Tail entity	Relationship	Sentence
1	E08	Broken needle	Fault_description	E08 represents a broken needle
2	E10	Check whether the manual reverse sewing switch is damaged	Fault_handling	The solution after the E10 fault code is to check whether the manual reverse sewing switch is damaged
3	E402	Pedal identification fault	Fault_causes	The possible cause of the E402 fault code is the pedial identification fault

In the original documents, relationships are annotated. For instance, in the sentence "E08 represents a broken needle", "E08" and "broken needle" are entities, and their relationship "fault_description" needs to be annotated accordingly.

Using the processed training set text, a relationship extraction model based on BiGRU-Attention is trained, with model parameter configurations outlined in Table 5. Subsequently, following the annotated results, test set texts are formatted as "head entity, tail entity, and containing sentence". The organized test set texts are then fed into the trained model to extract entity relationships, after which the results are compared. The evaluation metrics for the outcome show an accuracy rate of 93.34%, a recall rate of 89.41% and an F1 score of 91.75%.

Table 5 Relationship extraction model parameters

Parameter	Value
Learning rate	1×10^{-5}
Batch size	6
Iterative algebra	100

4 Conclusions and Future Work

By introducing and implementing the entity extraction model based on BERT-CRF and the relationship extraction model based on BiGRU-Attention, this paper provides an effective technical means for unstructured text information mining in the operation and maintenance management of sewing equipment, and provides a suitable supplementary case for the construction of KGs of sewing equipment.

Based on the sewing process corpus and service demand, combined with the industry knowledge structure characteristics and service system requirements, the design method of the sewing knowledge ontology model was studied. A total of 658 pieces of data were divided from sewing process documents and sewing equipment process manuals. According to the ontology and BIO three-dimensional annotation methods constructed, each text of the data was annotated, and a total of 423 entities were annotated, which had the characteristics of wide knowledge coverage and strong generalization ability, and realized the unified and standardized expression of unstructured knowledge.

Through the training of large-scale professional corpus, the above model can not only accurately identify and extract key entities and attributes in text, but also effectively capture complex relationships between entities. The experimental results show that compared with the existing deep learning algorithms, the model proposed in this paper can achieve a 20%–30% improvement in the recall rate and the accuracy rate, and effectively improve the speed and the accuracy of sewing equipment fault diagnosis.

The construction of the language model can accurately and quickly extract key entity information, accelerate equipment fault diagnosis and maintenance decision, and improve operation and maintenance efficiency. The application of advanced NLP technology to transform unstructured text into a searchable and reasonable knowledge base, effectively supports intelligent management and maintenance, and promotes the intelligent process of the industry. Deep mining of equipment knowledge and construction of KGs can effectively reduce cost and increase efficiency, improve product quality, and provide key technical support for the digital transformation of the manufacturing industry.

Future work will further optimize the performance of the model, expand the application range of the model, and explore more practical application scenarios. At the same time, we will also strive to integrate these technologies with other intelligent tools and platforms to provide more comprehensive and efficient solutions for the intelligent management of the operation and maintenance of sewing equipment and other manufacturing equipment.

References

[1] XU Y N, THOMASSEY S, ZENG X Y. AI for apparel manufacturing in big data era: a focus on cutting and sewing [M]//Artificial Intelligence for Fashion Industry in the Big Data Era. Singapore: Springer, 2018: 125-151.

[2] JIA J, YANG Q, FU H, et al. Construction of pre-trained language model and text semantic analysis based on power equipment big data [J]. *Chinese Journal of Electrical Engineering*, 2023, 43(3): 1027-1037. (in Chinese)

[3] TAO J Q, LI X Y, ZHENG P, et al. State-of-the-art and frontier of manufacturing KG application [J]. *Computer Integrated Manufacturing System*, 2022, 28(12): 3720-3736. (in Chinese)

[4] ZHANG D H, LIU Z Y, JIA W Q, et al. A review on knowledge graph and its application prospects to intelligent manufacturing [J]. *Journal of Mechanical Engineering*, 2021, 57(5): 90-113. (in Chinese)

[5] SHEN X W, LI X Y, ZHOU B, et al. Dynamic knowledge modeling and fusion method for custom apparel production process based on knowledge graph [J]. *Advanced Engineering Informatics*, 2023, 55: 101880.

[6] HU Y, DING Y, Xu F, et al. Knowledge recommendation system for human-robot collaborative disassembly using knowledge graph [C]//ASME 16th International Manufacturing Science and Engineering Conference. New York: ASME, 2021: 21-25.

[7] HEDBERG T D, MANAS B, CAMELIO J A. Using graphs to link data across the product lifecycle for enabling smart manufacturing digital threads [J]. *Journal of Computing and Information Science in Engineering*, 2020, 20(1): 011011.

[8] HUET A, SEGONDS F, PINQUIE R, et al. Context-aware cognitive design assistant: implementation and study of design rules recommendations [J]. *Advanced Engineering Informatics*, 2021, 50: 101419.

[9] HAO Q, HU S L. Smart design on the flexible gear based on knowledge graph [J]. *Journal of Physics: Conference Series*, 2021, 1885(5):

052021.

[10] XU D Z, JIN C, MA C, et al. Geographic named-entity recognition based on BERT-BiGRU-CRF and multi-head attention mechanism [J]. *Network Security and Data Governance*, 2023, 42(1): 169-173. (in Chinese)

[11] ZHANG Q, ZHAO G Y, SU Y, et al. Power text semantic recognition algorithm based on improved Bert-AutoML [J]. *Electronic Design Engineering*, 2024, 32(4): 43-46, 51. (in Chinese)

[12] GAO G Z, LI Y, HUA Y P, et al. Named entity recognition in oil and gas domain based on BERT-BiLSTM-CRF model [J]. *Journal of Changjiang University*, 2024, 21(1): 57-65. (in Chinese)

[13] DEVLIN J, CHANG M W, LEE K, et al. BERT: pre-training of deep bidirectional transformers for language understanding [C]// The North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Stroudsburg: ACL, 2019: 4171-4186.

[14] Li F, PAN H S, SHENG S X, et al. Image retrieval based on vision transformer and masked learning [J]. *Journal of Donghua University (English Edition)*, 2023, 40(5): 539-547.

[15] ZHOU H R, SONG H, GONG M M. Semantic path attention network based on heterogeneous graphs for natural language to SQL task [J]. *Journal of Donghua University (English Edition)*, 2023, 40(5): 531-538.

[16] ZHENG L R, XIAO X X, ZOU B J, et al. Named entity recognition for electronic medical record based on BERT [J]. *Computer and Modernization*, 2024, 1: 87-91. (in Chinese)

[17] YANG Y P, WANG S T. Study on malicious traffic classification algorithm based on CNN combined with BiGRU [J/OL]. *Computer Science*, 2024: 1-9 [2024-01-26]. https://kns.cnki.net/kcms/detail/50.1075.TP.20240124.1838.002.html. (in Chinese)

[18] LAN W W, XU W. Character-based neural networks for sentence pair modeling [C]// Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Stroudsburg: ACL, 2018: 157-162.

面向缝纫设备运维管理的语言模型构建方法研究

刘　冰[1]，刘　莹[1]，郑小虎[2*]，李靡晨[3]，杜思淇[4]

1. 杭州中服科创研究院有限公司，浙江 杭州 311103
2. 东华大学 人工智能研究院，上海 201620
3. 东华大学 信息科学与技术学院，上海 201620
4. 东华大学 机械工程学院，上海 201620

摘　要： 缝纫设备的智能运维与管理，关键在于解决非结构化文本的信息挖掘及语言模型构建问题。这对于加快设备缺陷和故障诊断速度、提高诊断准确性及实现设备检修的智能辅助决策，具有重要意义。该研究提出了通过基于 BERT 的条件随机场（bidirectional encoder representations from transformers-conditional random field，BERT-CRF）的实体抽取模型抽取关键实体信息，如设备名称、属性等，再通过基于双向门控循环单元注意力机制（bidirectional gated recurrent unit-attention，BiGRU-Attention）的关系抽取模型有效捕捉实体之间的语义关联，为缝纫设备知识图谱的构建提供支持。针对缝纫设备文本分析场景，模型在缝纫设备文本实体识别、信息抽取和故障诊断等任务场景进行了专门的训练和优化。与现有的深度学习算法相比，该研究所提方法在验证集和测试集上实现了 20%到 30%的性能提升，体现了其在召回率和精确度上的显著优势。缝纫设备知识的非结构化文本信息挖掘，可为平缝设备数据集成、设备故障运维、平缝工艺路线设计等方面的知识图谱构建提供参考。

关键词： 缝纫设备运行管理；语言模型；实体抽取；关系抽取；知识图谱