

Article A Text-Oriented Fault Diagnosis Method for Electromechanical Device Based on Belief Rule Base

Manlin Chen^{1,2}, Zhijie Zhou^{1,*}, Xiaoxia Han¹ and Zhichao Feng¹

- ¹ High-Tech Institute of Xi'an, Xi'an 710025, China; cmlwg@163.com (M.C.); wangwei@ccut.edu.cn (X.H.); 19bd04078@stu.hit.edu.cn (Z.F.)
- ² The School of Applied Technology, Changchun University of Technology, Changchun 130012, China
- Correspondence: zhouzj04@tsinghua.org.cn

Abstract: At present, quantitative data is often used for fault diagnosis of electromechanical devices, while qualitative data in the form of text is rarely used. In order to integrate qualitative data in the form of text and quantitative data in the fault diagnosis of an electromechanical device, a textoriented fault diagnosis method based on belief rule base (BRB) is proposed in this paper. Specifically, the key information of fault diagnosis is extracted from the text through natural language processing (NLP) and then converted into belief rules. Then, a rule supplement method is adopted to add the extracted belief rules to the BRB for the completion of the BRB construction. This method applies qualitative data in the form of text to the process of BRB construction, which is a new attempt at the BRB construction method. It not only solves the problem that BRB cannot use qualitative data in text form but also improves the modeling accuracy and data comprehensive processing ability of BRB. To verify the effectiveness of the algorithm, we designed an experiment of asynchronous motor fault diagnosis in the case study. The experimental result shows that the proposed method can use qualitative data in text form to construct BRB and effectively diagnose faults of asynchronous motors. The MSE of the proposed method is 0.0451, which is better than that of traditional BRB (0.1461), BP (0.0613), and SVR (0.0974) under the same experimental conditions.

Keywords: qualitative data; text form; belief rule base; NLP

MSC: 37M10

1. Introduction

The electromechanical device is an important part of all kinds of large equipment; its reliability is closely related to the safe operation of equipment, which is very important. Fault diagnosis is the main method to improve the safety of electromechanical devices [1–3]. At present, fault diagnosis methods can be divided into three categories [4]: (1) data-driven method [4,5], (2) model-based method, and (3) expert system. To effectively fault diagnosis, it is necessary to analyze the advantages and disadvantages of each method.

In general, data-driven and model-based methods have received attention in recent decades, with the characteristics of data processing and analysis of the model system to deal with faults. However, the data-driven method belongs to the "black box modeling" method; its internal calculation process and output lack interpretation [6,7]. The model-based methods require mathematical models designed according to the working principle of the object of study, which is difficult to implement for complex systems [8]. The expert system method can effectively combine model and data information to provide a more accurate fault diagnosis by comparing the advantages and disadvantages of the above two methods. The belief rule base (BRB) is similar to an expert system, which can use qualitative data and quantitative data to analyze quantitative data [9]. It is based on the traditional IF-Then rule, proposed by Yang et al. [6–8] in 2006. In recent years, BRB has been successfully applied in the field of fault diagnosis [9].



Citation: Chen, M.; Zhou, Z.; Han, X.; Feng, Z. A Text-Oriented Fault Diagnosis Method for Electromechanical Device Based on Belief Rule Base. *Mathematics* **2023**, *11*, 1814. https://doi.org/ 10.3390/math11081814

Academic Editor: João Paulo Costa

Received: 8 February 2023 Revised: 31 March 2023 Accepted: 4 April 2023 Published: 11 April 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). However, most of the current research on BRB fault diagnosis focuses on how to improve the diagnosis accuracy. There is a lack of research on how qualitative data in text form can be used in BRB. As a kind of widely used equipment, electromechanical device accumulates a lot of valuable qualitative data in the form of text in the working process, which can be used for fault diagnosis. However, due to the fact that BRB cannot use these data at present, it not only causes the waste of data but also directly affects the fault diagnosis effect of BRB and becomes the main factor restricting the further development of BRB. Therefore, how to apply qualitative data in text form to fault diagnosis of electromechanical equipment based on BRB has become an urgent problem to be solved.

Therefore, this paper proposes a text-oriented fault diagnosis method for electromechanical devices based on a belief rule base in order to improve the modeling accuracy of the fault diagnosis method based on BRB by using qualitative data in text form. For text information processing, Natural Language Processing (NLP) technology is a standard method [10–12]. It mainly studies how computers understand human language and perform relevant operations according to the results of their understanding [11–13]. At present, NLP has been used in many fields [11]. In this study, NLP is used for word segmentation of qualitative data in text form.

Through the use of the method proposed in this paper, the quantitative data in textual form is no longer wasted. The belief rules can be extracted from the text information related to the fault diagnosis of electromechanical devices. It improves the accuracy and interpretability of the initial parameters of the BRB and provides a guarantee for maintaining the interpretability of the parameter optimization results and preventing the optimization results from falling into the local optimum. In addition, how to reduce the influence of insufficient expert knowledge on the construction of BRB has always been the focus of research on the construction method of BRB. This method provides a way to solve these problems.

A case study at the end of this paper verifies the effectiveness of the proposed method. First, two confidence rules are extracted from the failure reports of electromechanical devices. Then, these two rules are complemented to an incomplete BRB to complete the construction of the BRB. Finally, the fault diagnosis of electromechanical devices is performed. Since belief rules are extracted from qualitative data in text form and no longer set by random function, the modeling accuracy of BRB is improved, and better fault diagnosis results are obtained.

2. Literature Review

Section 2.1 reviews the present studies on the data-driven method and model-based method. Section 2.2 reviews the current research status of BRB. So far, no studies have been conducted on the use of qualitative data in the form of text. Most of them focus on the BRB structure. Section 2.3 summarizes the current research status of NLP. So far, no combination of NLP and BRB has been found.

2.1. Data-Driven Method and Model-Based Method

Aiming at the data-driven method, Ikotun et al. [1] studied the single-wire grounding fault, dual-wire grounding fault, and three-phase grounding fault of an AC system inverter and their performance in a DC link and rectifier AC system and established a simulation model for high-voltage direct current (HVDC) single-pole system by using Matlab/Simulink software. Li et al. [3] proposed a data-driven optimal test selection design for fault detection and isolation. Hu et al. [4] proposed a new cause–effect relations method to determine the root cause of anomalies in industrial process monitoring of complex processes. Chen et al. [5] proposed a single-side neural network-aided canonical correlation analysis method for fault diagnosis of industrial systems. Although the above methods have achieved good results in fault diagnosis research, they can only use quantitative data. Those results in a waste of data and affect the effectiveness of fault diagnosis. In modelbased methods, Andrei S. Maliuk et al. [14] proposed a hybrid feature selection method based on Wrapper-WPT, used to improve the accuracy of early bearing fault diagnosis and achieved good results. Zhang et al. [15] proposed a method of rolling bearing fault diagnosis based on the threshold acquisition U-Net, which was successfully verified on the vibration signals of rolling bearings. Although the model-based method has good detection and diagnosis results, it is more difficult to accurately master the operation mechanism and model structure of the subject before use.

2.2. BRB

Due to its compelling combination of quantitative data and qualitative data, BRB has been successfully applied in equipment fault diagnosis, state assessment, and other fields [8,16,17]. Cheng et al. [18] proposed an effective fault diagnosis model based on BRB to solve the fault problem of the train running gear systems. Experiments showed that this method could locate faults more accurately. Feng et al. [19] put forward a BRB safety assessment method with attribute reliability, verified this method based on Liquefied Natural Gas (LNG) tanks and achieved good results. Tan et al. [20] proposed a collaborative distributed multiobjective method (CDMO-BRB) to optimize the structure and parameters of heterogeneous BRBs simultaneously. The method is validated in pipeline leakage detection, and the results show that CDMO-BRB has better performance than previous studies. Cao et al. [9] proposed a BRB-based health status assessment model in 2021, which successfully evaluated the operating status of aero-engines and proved the interpretability of BRB. Although the above studies have achieved good results, there is a lack of discussion on using qualitative data in text form. This calls for a new BRB capable of qualitative data processing using textual forms.

2.3. NLP

NLP can process text information and has functions such as word segmentation, semantic analysis, and syntactic analysis [10,13,21]. Zhao et al. [13] used the BoW model in NLP to count the fault-related words in the fault description and then sorted the statistical results to realize the extraction of fault features in mechanical fault information. Naseri Hossien et al. [21] used NLP to extract the pain which was summarized by doctors from clinical reports, and tested it by using the clinical words of patients with cancer bone metastasis, achieving good results and effectively verifying the ability of NLP to extract text information. Shi et al. [22] used NLP to extract information from FDA drug labels to enhance the product-specific guidance evaluation and achieved good results. Dastan Hussen Maulud et al. [23] explained the application of NLP in Internet searches and compared the performance of various application technologies in NLP. The result showed that all technologies reached a good level. These applications and studies show that NLP solves many text-processing problems in many fields, but the combination of NLP and BRB has not yet been discovered.

The above literature review shows that it is feasible to apply NLP to the fault diagnosis method based on BRB to achieve the purpose of using qualitative data in text form in fault diagnosis. Therefore, this paper proposes a text-oriented fault diagnosis method for electromechanical devices based on a belief rule base. This method uses NLP to extract the critical information describing the fault cause from qualitative data in text form and then convert it into belief rules. After that, the obtained belief rules are added to the BRB construction process. Finally, the BRB is adopted for fault diagnosis. This paper has two main innovations: (1) NLP is used to convert the fault description information in the text information into belief rules; (2) the constructed belief rules are added to the existing belief rules to complete the construction of BRB. The specific content and experimental verification of the method is explained in the following seven sections.

3. Proposed Approach

Focusing on the problem of electromechanical device fault diagnosis, the whole work of this paper is divided into seven steps, as shown in Figure 1. Starting from the fault diagnosis of an electromechanical device, this article first briefly introduces the main methods

of fault diagnosis. Then, the problem that qualitative text data cannot be used in fault diagnosis method based on BRB is found, and the solution to this problem is put forward. This method is divided into two parts: (1) extraction of belief rules and (2) supplementation of belief rules. Each part is introduced in detail. After that, the effectiveness of the method is illustrated by method verification and comparative study. Finally, the application of this method in the fault diagnosis of electromechanical devices is summarized and prospected.



Figure 1. Step of whole work.

4. Problem Formulation

Based on traditional IF-THEN rules, the BRB is an excellent method to solve the problem of the unified use of quantitative data and qualitative data [6,19,20]. Specifically, the IF part of each rule is composed of several attributes with weights, and the conclusion of the THEN part has a belief degree [6,23]. BRB is generally described as follows [23,24]:

 $R_i: IF(x_1 i s A_1^i) \cap (x_2 i s A_2^i) \dots (x_n i s A_n^i) THEN\{(D_1, \beta_{1,i}), (D_2, \beta_{2,i}), \dots, (D_m, \beta_{m,i})\}$ with the rule weight θ_i and the attribute weight $\delta_{1,i}, \delta_{2,i}, \dots, \delta_{n,i}$

where R_i represents the *i*th (i = 1, 2, ..., L) rule. x_j is the *j*th attribute (j = 1, 2, ..., n). A_j^i is the value of the *j*th attribute of R_i . D_k is the *k*th conclusion (k = 1, 2, ..., m). $\beta_{k,i}$ is the belief degree of the *k*th conclusion in R_i . If $\sum_{k=1}^{m} \beta_{k,i} = 1$, then R_i is complete; otherwise, R_i is incomplete. θ_i is the weight of R_i . $\delta_{j,i}$ is the weight of the *j*th premise property in R_i . The larger the weight value, the more important the content is [25].

Although BRB can effectively use quantitative and qualitative data, there is still no feasible method to use qualitative data in text. Therefore, finding a way to apply quality data in text form to BRB is urgent. In this regard, it is considered to use NLP technology to extract the information describing the mechanism of the device or analyze the fault cause from the text, then convert the information into belief rules and add them to the BRB to complete the construction of the BRB. To accomplish the above tasks, this paper needs to solve two fundamental problems:

Problem 1 The belief rule has four parts: the value of the attribute, the weight of the attribute, the weight of the belief rule, and the initial belief degree of each conclusion [6]. Additionally, the text used to supplement the expertise, which is made up of statements,

is the textual description. Therefore, the first problem to be solved in this paper is how to find the information that can be used to build belief rules from these text descriptions and convert the information into data that conforms to the format of belief rules. Problem 1 can be formulated as

$$R'_{u}(A'^{u},\delta'_{u},\beta'_{u},\theta'_{u}) = \Gamma(m)$$
⁽¹⁾

where *m* represents the text information. $\Gamma(\cdot)$ represents the transformation method adopted by the model. $R'_u(AI^u, \delta'_u, \beta'_u, \theta'_u)$ represents a belief rule obtained from the text ($0 \le u \le L$). AI^u and δ'_u represent the set of values and weights of each premise attribute in this rule, respectively. β'_u represents the belief set of each conclusion in the rule, and θ'_u represents the weight of the rule.

Problem 2 The belief rules R'_u obtained can be added to the existing BRB as the supplementary rule. However, the specific situation between the rule R'_u and the existing BRB needs to be considered in this supplement. In Formula (2), a rule supplement algorithm should be established to complete the BRB supplement operation.

$$BRB(R_i) = F(R'_u) \tag{2}$$

where $F(\cdot)$ is a nonlinear mapping for the adopted algorithm by the models and $BRB(R_i)$ represents the entirely reconstructed model.

To solve the above problems, this paper proposes a text-oriented fault diagnosis method for electromechanical devices based on BRB. The method is divided into two parts: (1) The NLP establishes the fault knowledge transformation algorithm and generates belief rules. (2) Establish the rule supplement algorithm, apply the obtained belief rules to the existing BRB, and complete the construction of the BRB. The concrete structure of the proposed method is shown in Figure 2.



Figure 2. Framework of method.

Before the method is executed, the expert needs to complete three tasks: (1) the establishment of a custom thesaurus, (2) the selection of a stop word list, and (3) the construction of an initial BRB. The custom thesaurus includes attribute vocabulary, attribute state vocabulary, fault vocabulary, and fault level vocabulary. It is possible that some belief rules in the initial BRB cannot determine the parameters and need to be supplemented.

In Part I of the method, a belief rule extraction algorithm based on NLP is established. This method first uses NLP to segment the text information of the device with the help of a custom thesaurus created in advance by experts and a selected stopword list. After that, attributes and attribute states are combined with the established combination rules to form a fault phenomenon phrase, and then faults and fault levels are combined to form a fault cause phrase. Then, the mapping relationship between the two phrases is established to complete the extraction of formal text rules. Finally, the attribute states and fault levels in the rules are converted into data that can be used to construct belief rules, completing the conversion of text to the confidence rule.

In Part II of the method, a belief rule supplement algorithm is established. This method can add the belief rules obtained in Part I to the initial BRB constructed by the expert in advance, supplement the belief rules that cannot determine parameters, or replace the existing belief rules at the choice of experts. If there are still rules in the initial BRB whose

6 of 25

parameters cannot be determined after rule supplementation, the parameters of these rules are assigned by random setting. Finally, the BRB construction is completed.

Below, the two algorithms are introduced in detail, respectively.

5. An Algorithm for Transforming Fault Knowledge into Belief Rules Based on NLP

To solve the problem 1 proposed in Section 3, this section proposes an algorithm based on NLP to transform fault knowledge into belief rules. First, using the NLP word segmentation tool and custom lexicon, belief rule information can be extracted from system fault diagnosis text in the Chinese context. Then, the information is formatted and converted into belief rules.

5.1. Rule Extraction

According to the structure of the belief rule, the information extracted from the text needs to be made up of two parts. The first part is the IF part of the rule, which describes the fault phenomenon, including the system attributes and the value of attributes. The second part is the THEN part of the rule, indicating the fault cause phenomenon, which is composed of the fault of the system and the belief degree of the fault. The key to extracting rules from text is to find the information that makes up the IF part and THEN parts. This requires the word segmentation function of NLP, combined with a custom thesaurus, to decompose the statements in the text into words [26]. Then, extract the key information from these words. Finally, this information is combined into a phrase describing the fault phenomenon and a phrase describing the fault cause phenomenon. Considering the complexity of Chinese and the possible irregular description of text information, this paper assumes that the description of fault information in the text to be processed is relatively accurate. The content of the custom thesaurus does not appear between the description of the phenomenon and the cause of the fault. Moreover, the description of the attribute state and fault level has been fully included in the attribute state vocabulary and fault level vocabulary. The specific process is described as follows:

Step 1: Build a custom thesaurus. According to device conditions, the expert builds custom vocabularies, including attribute vocabulary, $X(x) = \{x_1, x_2, ..., x_n\}$, attribute state vocabulary, $S(A) = \{A_1, A_2, ..., A_{\tilde{n}}\}$, fault vocabulary, $F(D) = \{D_1, D_2, ..., D_m\}$, and fault level vocabulary, $L(p) = \{p_1, p_2, ..., p_{\tilde{m}}\}$. Add a stop word list, $Stop(w) = \{w_1, w_2, ..., w_z\}$. Among them, the attribute vocabulary describes the possible conditions for indicator state and specifies the symbolic representation for each. The fault vocabulary describes various faults that can occur on a device. The failure vocabulary contains a description and a probabilistic representation of the probability of each failure. The stop-word list contains high-frequency words that need to be ignored, and it can be selected from several well-known stop-word lists in the field of NLP.

Step 2: Obtain the word segmentation set. The Chinese text is loaded, and then the word segmentation tool of NLP is used to perform a word segmentation operation on the text to obtain the word segmentation set composed of basic words and their corresponding parts of speech, $O(o) = \{o_1, o_2, \dots, o_m\}$.

Step 3: Obtain a collection of phrases to be grouped using a custom thesaurus. O(o) is traversed, and the useless words are found and removed from it by using the stop word table, namely O'(o'). Based on the attribute vocabulary, attribute state vocabulary, fault vocabulary, and fault level vocabulary, the corresponding basic words are found from O'(o') and recombined to obtain the fault symptom phrase and fault cause phrase and set labels p and f for them, respectively. Finally, put these phrases into the words set to be grouped $phrase_set(ph) = \{\{p: ph_1\}, \{f: ph_2\}, \{f: ph_3\}, \dots, \{p: ph_n\}\}$.

Step 4: Obtain the set of rules. The *phrase_set(ph)* is traversed, and the fault cause phrases corresponding to the fault phenomenon phrases are found by using the labels of each element in the *phrase_set(ph)*, and then these phrases are put into the final rule set.



The above process is shown in Figure 3.

Figure 3. Fault symptom and fault level extraction process.

The following uses the fault description "当电机转角过大, 电源电压高于阈值时, 很可能是因为电机卡死, 也有可能是由于电源短路. (When the motor Angle is larger and the power supply voltage is higher than the threshold, it is likely to be because the motor is stuck or the power supply is short-circuited.)" as an example to describe the fault symptom and fault cause extraction process.

Firstly, based on the knowledge of the device, the expert can define the attribute vocabulary X(x) and the fault vocabulary F(D), which is (" 电机转角" (motor rotation), " 电源电压" (supply voltage)) and (" 电机卡死" (motor stick), " 电源短路" (power supply circuit)). Then, according to the possible conditions of attributes, the state vocabulary of attributes S(A) is established, ((" 大 (large), 高 (high)": "H"), (" 小 (small), 低 (low)": "L")). According to the expression habit of fault possibility in the Chinese context, the vocabulary of fault level L(p) is established, ((" 极可能 (most probable)": 0.9), (" 很可能 (likely)": 0.7), (" 可能 (probable)": 0.4), (" 也许 (maybe)": 0.1)). The meanings of the above Chinese nouns are shown in Table 1.

Table 1. The meaning of Chin	lese nouns.
------------------------------	-------------

Chinese	English	
	motor rotation	
电源电压	supply voltage	
电机卡死	motor stick	
电源短路	power supply circuit	
大	large	
高	high	
小	small	
低	low	
极可能	most probable	
很可能	likely	
可能	probable	
也许	maybe	

Secondly, the word segmentation tool and Stop(w) are used for word segmentation operation to obtain O(o), which contains the basic word and its corresponding part of speech, as shown in Figure 4.



Figure 4. The base word set O(o).

Thirdly, the basic words are combined according to X(x) and F(D) to obtain the attribute phrase and the fault cause phrase, as shown in Figure 5.



Figure 5. The attribute phrase and fault cause phrase.

In the Chinese context, the commonly used grammatical structure is "attribute + state", and adjectives dominate the description of the state. Therefore, this can be taken as a condition. After finding an attribute, continue to look for an adjective appearing in S(A) to form a fault phenomenon phrase with the attribute. The commonly used grammatical structure for the fault cause is "level of failure + cause of failure", and the description word of the level of failure is an adverb. With this condition, after finding a fault cause, backtrack to the left to find an adverb that appears in L(p) to form a fault cause phrase with the fault cause. Put the successively obtained phenomenon phrase and cause phrase together with their labels "p" and "f " into phrase_set(ph), as shown in Figure 6.

Finally, *phrase_set(ph)* elements are traversed to establish the relationship between fault phenomenon and fault causes. Since the structure of the fault description statement is to introduce the phenomenon of the fault first and then explain the fault cause, the phenomenon phrase and cause phrase in the phrase set will be adjacent. Therefore, element search with "*p*" label can be carried out from *phrase_set(ph)*. When an element with "*f*" label is found, it means that the previously found element with "*p*" label matches this one with "*f*" label. It is put into *rule_set({p:x}, {f:y}) = {({p:x1}, {f:y1}), ({p:x2}, {f:y2}), ..., ({p:xn}, {f:yn})} as a found text rules, as shown in Table 2. "<i>p*" and "*f*" indicate fault phenomenon labels and fault cause labels, and "*x*" and "*y*" indicate fault symptoms and fault causes in a rule.



Figure 6. The set of phenomenon and cause phrases.

Table 2. *rule_set*({*p*:*x*}, {*f*:*y*}).

	р		f
<i>x</i> ₁	电机转角大, 电源电压高 (Larger motor Angle and higher power supply voltage)	y_1	很可能电机卡死, 可能电源短路 (It is possible that the motor is stuck or the power supply is short-circuited)

5.2. Conversion of Rules

The contents in the *rule_set* need to be converted to belief rules before they can be added to BRB. This conversion process extracts information from contents in text form, then converts it into parameter values of belief rules. According to the description of the rule R_i in Section 2, the parameters in a rule include the value A_j^i of the *j*th attribute x_j and the belief $\beta_{k,i}$ of the *k*th conclusion, the weight θ_i of R_i , and the weight $\delta_{j,i}$ of the *j*th attribute. Therefore, extracting these four values from the *rule_set* is the key to rule transformation. The specific steps are as follows:

Step 1: Calculate the attribute weight $\delta_{j,i}$. A full-text retrieval for the text information is performed to find the frequency of occurrence of each attribute in the attribute vocabulary and the frequency of occurrence of various faults in the fault vocabulary, denoted as AC_j (j = 1, 2, ..., n) and FC_k (k = 1, 2, ..., m). Then, calculate the proportion of each AC_j in the total number of attribute occurrences AT_j , and the proportion of each FC_k in the total number of faults occurrence k. The larger the values of AT_j and FT_k are, the more times attribute x_j and fault k are mentioned in the text, the more important x_j is, and the more frequently fault k occurs. Therefore, AT_j can be used as attribute weight, as shown in Formula (3):

$$\delta_{j,i} = AT_j = \frac{AC_j}{\sum\limits_{j=1}^n AC_j}$$
(3)

The calculation formula of AT_i is shown in Formula (4):

$$FT_j = \frac{FC_k}{\sum\limits_{k=1}^{m} FC_k}$$
(4)

Step 2: Compute the value A_j^i of the premise property x_j . Using RT_j as the standard, descending order is performed in the attributes of fp of each rule in $rule_set(\{p:x\}, \{f:y\})$ to unify the format of the IF part. The content of p label is traversed in each rule by using the content in the attribute vocabulary, and the attribute content is removed, and only adjectives describing the state are retained. Then, these state words are matched with the contents of the attribute state vocabulary to obtain the symbolic representation of each state.

Step 3: Calculate the *k*th conclusion, that is, the belief $\beta_{k,i}$ of the *k*th fault in the rule. With the FT_k obtained in Step 1 as the standard, the fault causes of *y* part of each rule in *rule_set*({*p*:*x*}, {*f*:*y*}) are sorted in descending order, and the format of the THEN part is unified. Then, the fault vocabulary is used to remove the fault words in each rule, and only the description words of fault level are reserved. For R_i , the probability values corresponding are found to the THEN part of the fault possibility descriptors in the fault level vocabulary $D_{k,i}$ (k = 1, 2, ..., m), and then Formula (5) is used to obtain the belief degree of this fault.

$$\beta_{k,i} = \frac{D_{k,i}}{\sum\limits_{k=1}^{m} D_{k,i}}$$
(5)

Step 4: Calculate rule weight θ_i . The initial value is $\tilde{\theta}_i$, as calculated by Formula (6).

$$\widetilde{\theta}_i = \frac{1}{\sum\limits_{i=1}^{L} i}$$
(6)

Later, look for the same rule in *rule_set*({*p*:*x*}, {*f*:*y*}). If the rules are the same, add the $\tilde{\theta}_i$ of these same rules to obtain the rule weight θ_i , as shown in Formula (7).

$$\theta_i = \tilde{\theta}_i + \tilde{\theta}_w, i = w, i, w = 1, 2, \dots, L$$
(7)

At this point, the four-parameter values of a belief rule are converted. Through the establishment of the above two parts of the model, the acquisition and transformation of text form rules are completed, respectively, and Problem 1 is solved.

6. A New Fault Diagnosis Method Based on BRB

To solve problem 2 raised in Section 3, a new fault diagnosis method based on BRB is proposed in this section. First, experts build the BRB with their expertise. Considering the possibility of insufficient expert knowledge, partial belief rules may not complete the construction; that is, the BRB of the construction is incomplete. At this point, the belief rules obtained in the previous section can be added to the incomplete BRB. Then, the BRB is optimized. Finally, the BRB is used for fault diagnosis.

6.1. Construction of BRB

Assuming that the set of all rules in BRB is $R = \{R_1, R_2, ..., R_i\}$, i = 1, 2, ..., L. Considering inadequate expert knowledge, only an incomplete BRB can be constructed. The rule set of incomplete BRB is $\tilde{R} = \{\tilde{R}_1, \tilde{R}_2, ..., \tilde{R}_i\}, i = 1, 2, ..., N$, and the belief rule set obtained from the text is $R' = \{R'_1, R'_2, ..., R'_{i'}\}, i' = 1, 2, ..., M$. The four situations are described as follows:

$$R = \tilde{R} \cup R', \, \tilde{R} \cap R' = \varnothing, \, L = N + M$$
(8)

$$R = \widetilde{R} \cup R', \, \widetilde{R} \cap R' = \emptyset, \, L < N + M$$
(9)

$$R \neq \widetilde{R} \cup R', \, \widetilde{R} \cap R' = \emptyset, \, L > N + M$$
⁽¹⁰⁾

$$R \neq \widetilde{R} \cup R', \, \widetilde{R} \cap R' = \emptyset, \, L < N + M$$
(11)

$$R \neq \widetilde{R} \cup R', \, \widetilde{R} \cap R' = \emptyset, \, L = N + M$$
(12)

$$R \neq \widetilde{R} \cup R', \, \widetilde{R} \cap R' \neq \emptyset, \, L > N + M$$
(13)

Among them, Formula (8) is the ideal case, in which R' exactly completes the missing part of R. Formula (9) indicates that R' completes not only the missing part of R, but also has some rules that are the same as some rules in \tilde{R} . Formula (10) shows that the rule in R' cannot complete the missing part of R, and they are the same as some rules in \tilde{R} . The above situation shows that the rule supplement operation is not a simple process of merging R' and \tilde{R} . It is necessary to find a reasonable way to complete the rule supplement under the condition of rational use of R'.

According to the cases listed in Formulas (8)–(13), the union operation of set R' and set \tilde{R} can be divided into two cases, namely $\tilde{R} \cap R' = \emptyset$ or $\tilde{R} \cap R' \neq \emptyset$. In these two cases, the rule \tilde{R} and R' can be compared on the condition that the attribute's value is the same to determine whether the same rule exists and perform different processing. The following contents are discussed separately.

- (1) When $\tilde{R} \cap R \neq = \emptyset$, it indicates that there is no overlap between the rules in \tilde{R} and the rules in R', that is, the rules in R' cannot be constructed due to insufficient expert knowledge. In this case, we can directly union R' and \tilde{R} .
- (2) When $R \cap R' \neq \emptyset$, it indicates that the rules in R and R' are repeated, that is, some rules in R' have been constructed by expert knowledge, and $R' \subseteq \tilde{R}$ may even appear. For duplicate rules, since they are the embodiment of expert knowledge, it is up to the expert to decide which rules to use. For the rule of no repetitions, the union of R' and \tilde{R} can be performed.

According to Formulas (10)–(13), after the union operation of R' and R is completed, the situation of $R \neq \tilde{R} \cup R'$ may occur, that is, there are still belief rules that cannot be constructed. For these rules, rule parameters can be set by random assignment to complete the construction of rules [16]. The above process is shown in Figure 7.



Figure 7. Rule supplement process.

The above operations complete the belief rules to supplement the incomplete BRB. The complete BRB is constructed, and Problem 2 is solved.

6.2. Reasoning Process of BRB

BRB is used for fault diagnosis. According to the basic structure of the BRB in Section 2, the sampling data e_j of attribute x_j needs to be converted into belief distribution $\alpha_{j,k}$ according to the pre-set attribute reference value $\gamma_{j,k}$, and the calculation formula is as follows [27]:

$$\alpha_{j,k} = \frac{\gamma_{j,k+1} - e_j}{\gamma_{j,k+1} - \gamma_{j,k}}, \gamma_{j,k} \le e_j \le \gamma_{j,k+1}$$

$$(14)$$

$$\alpha_{j,k+1} = 1 - \alpha_{j,k}, \gamma_{j,k} \le e_j \le \gamma_{j,k+1}$$
(15)

$$\alpha_{j,s} = 0, s = 1, 2, \dots, J_k, s \neq k, k+1$$
 (16)

The resulting $\alpha_{i,k}$ is used in the following formula [24]:

$$\omega_{i} = \frac{\theta_{i} \prod_{j=1}^{M} (\alpha_{j}^{i})^{\overline{\delta}_{j}}}{\sum\limits_{l=1}^{L} \theta_{l} \prod\limits_{j=1}^{M} (\alpha_{j}^{i})^{\overline{\delta}_{j}}}$$
(17)

where θ_i represents the weight of rule *i*th. α_j^i represents the belief distribution of the value of *j*th feature under rule *i*th, and $\overline{\delta}_j$ represents the weight of the feature. The numerator on the right-hand side represents the belief of the input value of each attribute in the *i*th rule considering the weight of rules, and then divides it with the belief of all rules represented by the denominator. ω_i on the left is the activation weight of the *i*th rule and $\omega_i \in [0, 1]$, $\sum_{i=1}^{L} \omega_i = 1$. When $\omega_i = 0$, it indicates that the rule is not activated; otherwise, the rule is triggered.

All activated rules of BRB are fused by ER method to obtain a belief distribution set for each conclusion, which can be expressed as [27]:

$$S(e) = \{ (D_m, \hat{\beta}_m), m = 1, 2, \dots, M \}$$
(18)

where *e* is a set of input vectors representing the input values of each feature. *S*(*e*) represents the belief set corresponding to each evaluation result obtained according to these inputs. $\hat{\beta}_m$ represents the belief degree relative to the evaluation result D_m , which is calculated by the ER analytic algorithm. The formula is as follows [27]:

$$\hat{\beta}_{m} = \frac{\mu \left[\prod_{i=1}^{L} \left(\omega_{i} \beta_{m,i} + 1 - \omega_{i} \sum_{m=1}^{M} \beta_{m,i} \right) - \prod_{i=1}^{L} \left(1 - \omega_{i} \sum_{m=1}^{M} \beta_{m,i} \right) \right]}{1 - \mu \left[\prod_{i=1}^{L} (1 - \omega_{i}) \right]}$$
(19)

$$\mu = \left[\sum_{m=1}^{M} \prod_{i=1}^{L} \left(\omega_{i}\beta_{m,i} + 1 - \omega_{i}\sum_{m=1}^{M} \beta_{m,i}\right) - (M-1)\prod_{i=1}^{L} \left(1 - \omega_{i}\sum_{m=1}^{M} \beta_{m,i}\right)\right]^{-1}$$
(20)

where ω_i is calculated from Formula (8) and represents the activation weight of the *i*th rule. $\beta_{m,i}$ represents the belief of the *m*th evaluation result in the *i*th rule, which is specified by the expert.

The obtained $\hat{\beta}_m$ is multiplied by the numerical representation $\mu(D_m)$ of the evaluation result D_m , that is, the utility of the evaluation result D_m is obtained. After summing

up each utility, the utility of a rule for all evaluation results is obtained, and the formula is as follows [27]:

$$\mu(S(e)) = \sum_{m=1}^{M} \mu(D_m) \hat{\beta}_m$$
(21)

6.3. Optimization of BRB

For BRB, its operation precision can be expressed by mean square error (MSE), and the formula is as follows [28]:

$$P = \frac{\sum_{i=1}^{L} \left[\mu(S(e)_i - y_i) \right]^2}{L}$$
(22)

where *P* represents the MSE. $\mu(S(e)_i)$ represents the utility of the input data e in BRB after processing in the *i*th rule, and y_i represents the actual result of the *i*th rule. *L* represents the number of rules in BRB. Therefore, the optimization goal of BRB is to make the error between the utility of each rule and the actual output value as small as possible. The objective function for parameter optimization can be established as follows [28]:

$$\begin{array}{ll} \min & \left\{ MSE(\delta_{j,i}, \beta_{k,i}, \theta_i) \right\} \\ s.t. & 0 \le \theta_i \le 1, i = 1, 2, \dots, L \\ & 0 \le \delta_{j,i} \le 1, j = 1, 2, \dots, n \\ & 0 \le \beta_{k,i} \le 1, k = 1, 2, \dots, m \\ & \sum_{k=1}^{M} \beta_{k,i} = 1 \end{array}$$

$$(23)$$

To solve the objective function shown in Formula (24), this paper uses the Projection covariance matrix adaptive evolution strategy (P-CMA-ES) as the optimization algorithm, which can solve nonlinear, non-convex real-value continuous optimization problems and is suitable for parameter optimization under various constraints of the objective function [29].

6.4. BRB Fault Diagnosis Algorithm

With the text information as the source of expert knowledge, the NLP technology is used to extract and transform rules. Then, the obtained belief rules are added to the incomplete BRB, and the BRB construction is completed. This process, shown in Figure 8, consists of the following steps.

Step 1: According to the research object, experts establish a custom thesaurus, including attribute vocabulary, attribute state vocabulary, failure vocabulary, failure level vocabulary, and stop word list.

Step 2: According to the created custom thesaurus, using the NLP, the loaded text information is processed by word segmentation, and the belief rule set $rule_set({p:fp}, {f:fa})$ in text form is obtained.

Step 3: Formulas (1)–(5) are used to format belief rules set $rule_set(\{p:fp\}, \{f:fa\})$ in text form and obtain BRB supplementary rule set R'.

Step 4: Experts construct the BRB and obtain an incomplete belief rule set \hat{R} in the case of insufficient expert knowledge.

Step 5: According to $R \cap R' = \emptyset$ or $R \cap R' \neq \emptyset$, the union operation of R' and R is carried out.

Step 6: If $R \neq R \cup R'$ still exists after the union operation, the rest rules are constructed by random assignment.

Step 7: The BRB is optimized by P-CMA-ES.

Step 8: Fault diagnosis is performed using the optimized BRB.



Figure 8. BRB construction model.

7. Case Study

To verify the validity of the above methods, this paper took a certain asynchronous motor as the research object. There are three reasons for choosing asynchronous motors as the research object. (1) Asynchronous motor is widely used and is a typical representative of an electromechanical device, which is relatively consistent with the research background of this article. (2) The design and manufacture of the asynchronous motor is mature, and the qualitative data in the form of text related to it is relatively easy to collect, which can provide data support for the experiment. (3) Many people have tried a variety of methods for asynchronous motor fault diagnosis, but it is rare to apply qualitative data in the form of text to the fault diagnosis process. The qualitative data is processed from the Chinese fault diagnosis report of asynchronous motor, and the belief rules of BRB are obtained after extraction and conversion. It is assumed that the expert knowledge is insufficient and can only construct an incomplete BRB. At this point, the obtained BRB supplementary rule can be added to the incomplete BRB. Finally, the BRB that completes the supplement is used for the fault diagnosis operation of the asynchronous motor.

7.1. Numerical Analysis

An asynchronous motor is the key part of the mechanical and electrical system, which has the characteristics of large amounts, easy wear, and various faults [30–32]. The fault diagnosis of its operation can be carried out by two attributes: motor angle and power supply voltage [33]. The motor rotation angle is collected by a rotary encoder, and the power supply voltage is obtained by a graphic recording voltmeter. A total of 72 samples are obtained. Among them, there are four operating states of the motor, which include 23 data of power short circuit (D1), 27 data of motor stuck (D2), 10 data of power disconnect (D3), and 12 data of normal (D4). It is shown in Figure 9.



Figure 9. Data graph.

The blue circles in the Figure 9 represent the values of the sampled data. Take half of each state data as the training set and the rest as the test set, as shown in Table 3.

Table 3. Experimental data.

Category	D1	D2	D3	D4
Train	12	14	5	6
Test	11	13	5	6

When the quantitative data scale is large, data-driven fault diagnosis methods can be used [34]. However, the experimental data scale is small, and the data-driven fault diagnosis method is no longer applicable [35]. However, because BRB can solve small sample data, it can be used for asynchronous motor fault diagnosis operations with fewer experimental data.

According to the working mechanism of a three-phase asynchronous motor and the obtained data, the reference levels of the motor rotation angle and the power supply voltage can be set as High (H), Low (L), and Normal (N). The reference values are shown in Table 4.

Table 4. Attribute reference value.

Attribute	Н	L	Ν
motor rotation angle	285	247	270
power supply voltage	400	300	380

According to the above analysis, the custom vocabularies are established, including attribute vocabulary X(x), attribute state vocabulary S(A), fault vocabulary F(D), and fault level vocabulary L(p). The specific contents are listed in Table 5.

In the Chinese context, many organizations have introduced stop-word lists. This example uses the Chinese stop word table presented by the Harbin Institute of Technology, which is rich in content and has 1893 stop words, which can meet the needs of most word segmentation operations [36].

For Chinese text word segmentation, NLP word segmentation technology is needed to complete the operation. At present, there are three main methods for Chinese word segmentation: rule word segmentation, statistical word segmentation, and mixed word segmentation (the combination of rules and statistics) [36–38]. Among them, regular word segmentation operation. It is relatively simple and efficient, but it is difficult to process the new words that do not appear in the custom thesaurus [36,37]. On the other hand, statistical word segmentation is a method that emerged after the rise of machine learning technology. It divides statements by establishing a word segmentation model and then calcu-

lates the probability of the segmentation results by statistical methods (such as the hidden Markov model, conditional random field, etc.) to obtain the word segmentation method with the maximum probability. Although this word segmentation method can better deal with the emergence of new words, the model needs training and is more dependent on the corpus quality during training [38,39]. Mixed word segmentation is a combination of regular word segmentation and statistical word segmentation. Therefore, it is better to accomplish word segmentation through the self-defined thesaurus and word segmentation training model. Thus, the current mainstream Chinese word segmentation engines, such as Jieba, Paoding, SnowNLP, etc., adopt the mixed word segmentation method [39]. In this example, the Jieba word segmentation engine is selected as the word segmentation tool to perform word segmentation operations on the fault diagnosis report of the asynchronous motor.

Table 5. Custom vocabularies.

Vocabulary	Content
X(x)	电机转角, 电源电压 (motor angle, power supply voltage)
S(A)	(过大 (larger), 高于 (higher)): H,
$\mathcal{S}(A)$	(过小 (too small), 低于 (lower)): L, (正常 (normal)): N
F(D)	电机卡死 (motor stuck), 电源短路 (power short circuit),
	电源断路 (power break), 正常 (normal)
I(n)	(极可能 (most probable):0.9, 非常可能 (very probable): 0.8, 很可能 (likely): 0.7,
L(p)	可能 (probable): 0.4, 也许 (maybe): 0.2)

7.2. Rule Extraction and Transformation

First, the fault diagnosis report of the asynchronous motor is analyzed using the custom thesaurus and the stop word table using the Jieba word segmentation engine. Then, the rule set is obtained in Table 6.

Table 6. The rules set.

_ <i>p</i>	f
电机转角大,电源电压高	很可能电机卡死,可能电源短路
(Larger motor angle, high power supply	(It is possible that the motor is stuck or the
voltage)	power supply is short-circuited)
电机转角大 (Larger motor angle)	极可能电机卡死 (Most likely, the motor is stuck)

The times of occurrence of each premise attribute in the fault diagnosis report are counted. The results are listed in Table 7.

Table 7. The times of occurrences of each attribute.

Attribute	Times
电机转角 (motor rotation angle)	38
电源电压 (power supply voltage)	73

Formula (1) is used to calculate the weight of each premise attribute as follows:

$$\delta_1 = \frac{38}{38 + 73} = 0.34 \ \delta_2 = \frac{73}{38 + 73} = 0.66$$

Then, according to the setting of the attribute state value in S(A), the attribute state value of the two rules in the rule set is "HH" and "H", respectively. Since rule 2 does not state the voltage of the supply, it is considered to be in a normal state, and its state value is "HN". In the fault diagnosis report, statistics are made on the occurrence times of each fault. The results are shown in Table 8 below:

Table 8. The times of failures.

Fault	Times	
电机卡死 (motor is stuck)	9	
电源短路 (power supply circuit)	12	
电源断路 (power disconnect)	7	

Formula (2) is used to calculate the frequency of each failure, which is as follows:

$$FT_1 = \frac{9}{9+12+7} = 0.32 \ FT_2 = \frac{12}{9+12+7} = 0.42 \ FT_3 = \frac{7}{9+12+7} = 0.25$$

Sort the parts of the two rules in the rule set according to the value of FT_k , and the results are shown in Table 9:

Table 9. Sorted rule set.

	p		f
<i>x</i> ₁	电机转角大, 电源电压高 (Larger motor angle, high power supply voltage)	y_1	可能电源短路, 很可能电机卡死 (The power supply may be short-circuited, and the motor may be stuck)
<i>x</i> ₂	电机转角大 (Larger motor angle)	<i>y</i> ₂	极可能电机卡死极可能电机卡死 (Most likely, the motor is stuck)

Then, according to the setting of various fault possibilities in L(p), Formula (3) is used to calculate the belief degree of known conclusions in each rule, which can be written as follows:

$$\beta_{1,1} = \frac{0.4}{0.4 + 0.9} = 0.3 \ \beta_{2,1} = \frac{0.9}{0.4 + 0.9} = 0.7 \ \beta_{1,2} = \frac{0.9}{0.9} = 1$$

For the conclusion without explanation, the belief can be set to 0.

Finally, according to Formulas (4) and (5), the weights of the two rules are 0.5 and 0.5, respectively. At this point, for the two rules extracted from the text, their transformed contents are listed in Table 10.

Table 10. Values of rule parameters after transformation.

	Attribute Values	Attribute Weight	The Belief	Rule Weight
R_1	HH	0.34, 0.66	0.3, 0.7, 0, 0	0.5
<i>R</i> ₂	NH	0.34, 0.66	0, 1, 0, 0	0.5

7.3. Added BRB Supplementary Rules

Based on the analysis of the premise properties and running state of the three-phase asynchronous motor, the ith rule of BRB is as follows:

 $\begin{aligned} R_i: IF \left(\text{motor rotation angle is } A_1^i \right) \cap \left(\text{power supply voltage is } A_2^i \right) \text{ THEN} \\ \left\{ \left(D_1, \beta_{1,i} \right), \left(D_2, \beta_{2,i} \right), \left(D_3, \beta_{3,i} \right), \left(D_4, \beta_{4,i} \right) \right\} \text{ with the rule weight } \theta_i \text{ and the attribute weight } \delta_{1,i}, \delta_{2,i} \end{aligned}$

Considering the lack of expert knowledge and according to the different values of the premise attribute of a three-phase asynchronous motor, BRB construction is carried out for the four conclusions. The premise attribute value is 1, and the weight of each rule is 0.8. Therefore, N/A indicates that it cannot be determined, as listed in Table 11.

	Attri	Attributes		Consequent			
Kule No.	Moto Rotation Angle	Power Supply Voltage	D ₁	D_2	D ₃	D_4	- Kule weight
1	Н	Н	N/A	N/A	N/A	N/A	
2	H(1)	L(1)	0.1	0.5	0.4	0	0.8
3	Н	N	N/A	N/A	N/A	N/A	
4	N(1)	H(1)	0.8	0.1	0.1	0	0.8
5	N(1)	L(1)	0.3	0.1	0.6	0	0.8
6	N(1)	N(1)	0	0	0	1	0.8
7	L	Н	N/A	N/A	N/A	N/A	
8	L(1)	L(1)	0.1	0	0.9	0	0.8
9	L	Ν	N/A	N/A	N/A	N/A	

Table 11. Incomplete BRB.

As can be seen from Table 11, there are four rules which cannot be constructed by expert knowledge. Among them, rules 1 and 3 are consistent with the BRB supplementary rules shown in Table 9 and can be added directly to the BRB. For rules 7 and 9, there is no corresponding supplementary rule, which needs to be set randomly. The conclusions in the two rules are set to 0.25, the weight of the rule is set to 1, and the weight of the attribute is set to 1. At this point, the situation of each rule in BRB is listed in Table 12.

Table 12. BRB for adding supplementary rules.

D 1 M	Attributes		Consequent				D.1. 1. 147 1. 1
Kule No.	Moto Rotation Angle	Power Supply Voltage	D1	D ₂	D ₃	D_4	- Kule weight
1	H(0.34)	H(0.66)	0.3	0.7	0	0	0.5
2	H(1)	L(1)	0.1	0.5	0.4	0	0.8
3	H(0.34)	N(0.66)	0	1	0	0	0.5
4	N(1)	H(1)	0.8	0.1	0.1	0	0.8
5	N(1)	L(1)	0.3	0.1	0.6	0	0.8
6	N(1)	N(1)	0	0	0	1	0.8
7	L(1)	H(1)	0.25	0.25	0.25	0.25	1
8	L(1)	L(1)	0.1	0	0.9	0	0.8
9	L(1)	N(1)	0.25	0.25	0.25	0.25	1

P-CMA-ES algorithm is used for optimization. The population size is set to 40, and the number of evolutions is set to 300. The optimized BRB is listed in Table 13.

Table 13. Optimized BRB.

D 1 M	Attributes		Consequent				
Kule No.	Moto Rotation Angle	Power Supply Voltage	D ₁	D ₂	D ₃	D ₄	- Kule weight
1	H(0.8)	H(1)	1	0	0	0	0.98
2	H(0.8)	L(1)	0.9	0.1	0	0	0.60
3	H(0.8)	N(1)	0	0	1	0	0.75
4	N(0.8)	H(1)	0	0.9	0.1	0	0.58
5	N(0.8)	L(1)	0	0	0	1	0.42
6	N(0.8)	N(1)	0	0.1	0.8	0.1	0.60
7	L(0.8)	H(1)	0	1	0	0	0.85
8	L(0.8)	L(1)	0	0.2	0.7	0.1	0.65
9	L(0.8)	N(1)	0.1	0	0.8	0.1	0.39

The optimized BRB is verified by using the data of the test set, and the fault diagnosis results of the operation of the asynchronous motor are obtained, as shown in Figure 10.



Figure 10. Fault diagnosis results.

It can be seen from Figure 10 that the optimized BRB can accurately distinguish the four operating states, MSE = 0.0451. In the four kinds of status, the belief distribution of each sample is shown in Figure 11. It can be found that the belief degree of many samples in their actual status does not reach 1, which explains why the diagnosis results of the samples in Figure 10 have some deviations from the actual results.



Figure 11. Belief distribution of diagnostic results of each sample.

7.4. Comparative Study

(1) Comparison with traditional BRB. In the traditional BRB building process, if there is a rule that cannot be built, then the random assignment is used to specify parameters [5]. In this example, four rules cannot be constructed by expert knowledge. Therefore, the belief distribution of each of the four rules can be divided equally, and both attribute weight and rule weight can be set to 0.25, as listed in Table 14.

D.I.N.	Attributes		Consequent				- Bula Walaht
Kule No.	Moto Rotation Angle	Power Supply Voltage	D1	D ₂	D ₃	D_4	- Kule weight
1	H(1)	H(1)	0.25	0.25	0.25	0.25	1
2	H(1)	L(1)	0.1	0.5	0.4	0	0.8
3	H(1)	N(1)	0.25	0.25	0.25	0.25	1
4	N(1)	H(1)	0.8	0.1	0.1	0	0.8
5	N(1)	L(1)	0.3	0.1	0.6	0	0.8
6	N(1)	N(1)	0	0	0	1	0.8
7	L(1)	H(1)	0.25	0.25	0.25	0.25	1
8	L(1)	L(1)	0.1	0	0.9	0	0.8
9	L(1)	N(1)	0.25	0.25	0.25	0.25	1

Table 14. Traditional BRB.

P-CMA-ES is used to optimize this BRB, and the population size and optimization times are still set as 40 and 300. The optimized results are shown in Table 15.

Table 15. Optimized traditional BRB.

D 1 M	Attri	Attributes		Consequent			
Kule No.	Moto Rotation Angle	Power Supply Voltage	D1	D2	D3	D4	- Kule weight
1	H(1)	H(0.9)	1	0	0	0	0.99
2	H(1)	L(0.9)	0.4	0.2	0.1	0.3	0.01
3	H(1)	N(0.9)	0	0	1	0	0.53
4	N(1)	H(0.9)	0	1	0	0	0.22
5	N(1)	L(0.9)	0	0	0	1	0.02
6	N(1)	N(0.9)	0	0	0.8	0.2	0.01
7	L(1)	H(0.9)	0	1	0	0	0.95
8	L(1)	L(0.9)	0.3	0	0.6	0.1	0.05
9	L(1)	N(0.9)	0.2	0.1	0.7	0	0.87

The optimized BRB is used to analyze the test set again, and the fault diagnosis result of the fault operation of the three-phase asynchronous motor is obtained, as shown in Figure 12.



Figure 12. Fault diagnosis results of traditional BRB.

It can be seen from Figure 12 that there is a certain deviation between the traditional BRB's diagnosis of the running status and the actual status, and there is a certain gap between the diagnosis result shown in Figure 10. This is because BRB is an expert system whose initial parameter settings rely heavily on expert knowledge. In this case, four of the nine rules cannot be constructed by expert knowledge, and nearly half of the rules are constructed by random assignment, resulting in too little expert knowledge in BRB, which inevitably affects the operation effect of BRB. From this point, the importance of the method proposed in this paper to build the BRB when insufficient expert knowledge can be demonstrated.

8. Discussion

In order to further verify the effectiveness and stability of the proposed method (BRB1), this paper conducted a comparative study with traditional BRB (BRB2), BP (Back Propagation neural network), and SVR (Support Regression Vector).

Use the training set data in Table 3 to train BRB2, BP, and SVR, respectively. After that, the trained model is tested using the test set data in Table 3. Compare the test results of BRB1, BRB2, BP, and SVR, as shown in Table 16 below.

Table 16. Comparison of test results of various methods.

	BRB1	BRB2	BP	SVR	
MSE	0.0451	0.1461	0.0613	0.0974	

From Table 16, it can be seen that the accuracy of BRB1 is the highest. Due to the random setting of four belief rules, the accuracy of BRB2 is lower than that of BP and SVR. However, the diagnostic processes of BP and SVR lack interpretability, which is a defect of their own.

Following is cross-validation to continue the comparative study of BRB1, BRB2, BP, and SVR. The 72 samples are divided into 6 groups, with 12 samples in each group, as shown in Table 17.

Group	D1	D2	D3	D4	
1	4	5	1	2	
2	4	4	2	2	
3	4	4	2	2	
4	4	4	2	2	
5	4	5	1	2	
6	3	5	2	2	

Table 17. Sample grouping.

Then, six groups of training sets and test sets are established according to the proportion of training sets and test sets of 50%, respectively, as shown in Table 18.

Table 18. A training set and test set	et.
---------------------------------------	-----

Train	Test	
1, 2, 3	4, 5, 6	
2, 3, 4	5, 6, 1	
3, 4, 5	6, 1, 2	
4, 5, 6	1, 2, 3	
5, 6, 1	2, 3, 4	

BRB1, BRB2, BP, and SVR are used for validation according to the six groups, and the average MSE of the six validations is calculated. According to this process, the proportion

of the training set and test set is adjusted to 40%:60% and 30%:70%. After that, six times of verification are carried out, respectively, to calculate the average MSE. The results of the above process are shown in Table 19.

Training Set Ratio	BRB1	BRB2	BP	SVR
50%	0.0671	0.1833	0.0877	0.1025
40%	0.0839	0.1969	0.1731	0.2421
30%	0.1033	0.2010	0.4112	0.5633
average	0.0848	0.1937	0.2273	0.3026

Table 19. Comparison results of different methods.

It can be seen from Table 19 that the average MSE of the method proposed in this paper (BRB1) is lower than that of the other three models in the case of three different proportions of training sets. Among them, the average MSE of BRB2 is higher than that of BRB1 due to the random set of four rules in traditional BRB (BRB2). BP and SVR have commonly used machine learning algorithms. Since they need the support of large data samples, their MSE is higher than BRB1, with only 72 data samples. In addition, the BRB1 also shows high stability, as shown in Figure 13.



Figure 13. Comparison of stability of the four models.

As shown in Figure 13, the rectangular box of BRB1 is the narrowest, and the median line is basically in the middle, indicating that the MSE of each diagnosis result in BRB1 is relatively concentrated and stable. SVR is also more stable, but it is less accurate. In summary, the above studies demonstrate the effectiveness and stability of the proposed method.

The above comparative study shows that it is feasible and efficient to extract belief rules from qualitative data in text form and apply them to the construction of BRB. This indicates that this method plays an important role in improving data utilization and improving BRB modeling accuracy. Moreover, by processing qualitative data in text form, the construction of BRB can become faster and more convenient. In the future, the role of experts may shift from rule-making to rule review, ultimately making it possible to build BRB automatically.

NLP is a specialized field of study. The method proposed in this paper is the first combination of NLP and BRB and does not involve more complex functional applications of NLP, such as sentiment analysis functions. If more NLP functions are added to the method, the processing ability of qualitative data in text form will be further enhanced, which will provide more powerful support for the construction of BRB.

9. Conclusions

Aiming at the problem that the fault diagnosis method for electromechanical devices based on belief rule base (BRB) cannot use qualitative data in the form of text, a textoriented fault diagnosis method for electromechanical devices based on BRB is proposed in this paper. Specifically, the key fault information conversion algorithm based on NLP is used to obtain belief rules from qualitative data in text form. By using the belief rule supplement algorithm, the obtained belief rules are added to the BRB to complete the construction of the BRB. Adding qualitative data in the form of text into the process of BRB construction is an innovation to the method of BRB construction. It not only brings the qualitative data in text form into the data range that BRB can process but also improves the modeling accuracy of BRB. Compared with the previous method of improving the accuracy of BRB modeling by updating the optimization algorithm, the proposed method is more intuitive and efficient. Through a case study, it is verified that the proposed method can extract belief rules from the fault diagnosis report of the induction motor and apply these rules to the construction of BRB. The case shows that the MES of fault diagnosis results is 0.0451. In subsequent comparative studies, the MSE of traditional BRB, BP, and SVR are 0.1461, 0.0613, and 0.0974, respectively, which are all lower than the method proposed in this paper. This indicates that the proposed method can be more effectively applied to the fault diagnosis of electromechanical devices and provides a new choice for fault diagnosis in practice.

Although the above methods have achieved good results, their ability to deal with more complex text information still has some limitations. It can be considered to add a semantic analysis function in the subsequent research to solve the problem. In addition, the proposed method can only deal with Chinese, and how to extract belief rules for qualitative data in other language text forms, such as English, has not been studied and analyzed. This is the direction that should be further studied in the future.

Author Contributions: Conceptualization, M.C. and Z.Z.; methodology, Z.Z.; validation, X.H.; formal analysis, M.C.; investigation, Z.F.; resources, Z.Z.; data curation, M.C.; writing—original draft preparation, M.C.; writing—review and editing, M.C.; visualization, M.C.; supervision, Z.Z.; project administration, Z.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Natural Science Foundation of China (Nos. 61773388, 61751304, 61833016, 61702142, U1811264, and 61966009), the Shaanxi Outstanding Youth Science Foundation, China (No. 2020JC-34).

Data Availability Statement: Not applicable.

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

References

- Ikotun, O.; Agyekum, E.B.; Ahmed, E.M.; Kamel, S. Using Matlab/Simulink Software Package to Investigate Fault Behaviors in HVDC System. *Mathematics* 2022, 10, 3014. [CrossRef]
- Chen, H.T.; Liu, Z.G.; Alippi, C.; Huang, B.; Liu, D.R. Explainable Intelligent Fault Diagnosis for Nonlinear Dynamic Systems: From Unsupervised to Supervised Learning. *IEEE Trans. Neural Netw. Learn. Syst.* 2022, 1–14. [CrossRef] [PubMed]
- Li, Y.; Chen, H.T.; Lu, N.Y.; Jiang, B.; Zio, E. Data-Driven Optimal Test Selection Design for Fault Detection and Isolation Based on CCVKL Method and PSO. *IEEE Trans. Instrum. Meas.* 2022, 71, 3512310. [CrossRef]
- Hu, W.; Wang, J.; Yang, F.; Han, B.; Wang, Z. Analysis of time-varying cause-effect relations based on qualitative trends and change amplitudes. *Comput. Chem. Eng.* 2022, 162, 107813. [CrossRef]
- 5. Chen, H.T.; Chen, Z.W.; Chai, Z.; Jiang, B.; Huang, B. A single-side neural network-aided canonical correlation analysis with applications to fault diagnosis. *IEEE Trans. Cybern.* **2022**, *52*, 9454–9466. [CrossRef]
- 6. Zhou, Z.J.; Hu, G.Y.; Hu, C.H. A survey of belief rule base expert systems. *IEEE Trans. Syst.* 2019, 51, 4944–4958. [CrossRef]
- Zhu, W.; Chang, L.; Sun, J. Parallel multipopulation optimization for belief rule base learning. *Inf. Sci.* 2020, 556, 436–458. [CrossRef]
- Liu, Z.Z.; Xiao, M.Q.; Zhu, H.Z. Fault diagnosis of missile refrigeration system based on the belief rule base. J. Phys. Conf. Ser. 2020, 1507, 082023–082030. [CrossRef]
- Zhou, Z.J.; Cao, Y.; Hu, G.Y.; Zhang, Y.M.; Tang, S.W.; Chen, Y. New health-state assessment model based on belief rule base with interpretability. *Sci. China Inf. Sci.* 2021, 64, 72214–72221. [CrossRef]

- 10. Zhang, C.; Lu, K. Knowledge Graph Completion Algorithm Based on Probabilistic Fuzzy Information Aggregation and Natural Language Processing Technology. *Mathematics* **2022**, *10*, 4578. [CrossRef]
- Guo, L.; Yan, F.; Li, T. An automatic method for constructing machining process knowledge base from knowledge graph. *Robot. Comput. Integr. Manuf.* 2022, 73, 102222. [CrossRef]
- 12. Bolívar, S.; Nieto-Reyes, A.; Rogers, H.L. Statistical Depth for Text Data: An Application to the Classification of Healthcare Data. *Mathematics* **2023**, *11*, 228. [CrossRef]
- 13. Zhao, D.F.; Liu, S.L.; Miao, Z.H.; Zhang, H.L.; Wei, Y.; Xiao, S.G. A Novel Feature Extraction Approach for Mechanical Fault Diagnosis Based on ESAX and BoW Model. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 3516011. [CrossRef]
- Maliuk, A.S.; Ahmad, Z.; Kim, J.-M. Hybrid Feature Selection Framework for Bearing Fault Diagnosis Based on Wrapper-WPT. Machines 2022, 10, 1204. [CrossRef]
- Zhang, D.; Zhang, L.; Zhang, N.; Yang, S.; Zhang, Y. Early Fault Diagnosis of Rolling Bearing Based on Threshold Acquisition U-Net. *Machines* 2023, 11, 119. [CrossRef]
- Chen, K.; Chen, S. Research on Fault Diagnosis Strategy of Dual Three-Phase Permanent Magnet Synchronous Motor Drive System Based on Current Residual Errors, presented at RICAI 2020. In Proceedings of the 2020 2nd International Conference on Robotics. Intelligent Control and Artificial Intelligence, Shanghai, China, 2 October 2020; pp. 409–414.
- Cheng, X.; Qian, G.; He, W.; Zhou, G. A Liquid Launch Vehicle Safety Assessment Model Based on Semi-Quantitative Interval Belief Rule Base. *Mathematics* 2022, 10, 4772. [CrossRef]
- Cheng, C.; Wang, J.H.; Zhou, Z.J.; Teng, W.X.; Sun, Z.B.; Zhang, B.C. A BRB-Based Effective Fault Diagnosis Model for High-Speed Trains Running Gear Systems. *IEEE Trans. Intell. Transp. Syst.* 2022, 23, 110–121. [CrossRef]
- Feng, Z.C.; He, W.; Zhou, Z.J. A New Safety Assessment Method Based on Belief Rule Base With Attribute Reliability. *IEEE/CAA* J. Autom. Sin 2020, 8, 1774–1785. [CrossRef]
- Tan, X.; Chang, L.L.; Chen, Y.W.; Hao, Z.Y.; Wu, G.H. Cooperative and Distributed Multiobjective Optimization for Heterogeneous Belief Rule Base. *IEEE Syst. J.* 2021, 16, 777–788. [CrossRef]
- Naseri, H. Feasibility of Using Natural Language Processing to Extract Cancer Pain Score from Clinical Notes. *Radiother. Oncol.* 2019, 139, S65. [CrossRef]
- Xuan, Q.A.; Jl, A.; Yw, A. Natural language processing was effective in assisting rapid title and abstract screening when updating systematic reviews—ScienceDirect. J. Clin. Epidemiol. 2021, 133, 121–129.
- Dastan, H.M.; Siddeeq, Y.A.; Naaman, O. Review on Natural Language Processing Based on Different Techniques. Asian J. Res. Comput. Sci. 2021, 10, 1–17.
- 24. Feng, Z.C.; Zhou, Z.J.; Hu, C.H.; Chang, L.L.; Hu, G.Y.; Zhao, F.J. A new belief rule base model with attribute reliability. *IEEE Trans. Fuzzy Syst.* 2019, 27, 903–916. [CrossRef]
- 25. Zhou, M.; Chen, Y.W.; Liu, X.B. Weight assignment method for multiple attribute decision making with dissimilarity and conflict of belief distributions. *Comput. Ind. Eng.* **2020**, *147*, 106648. [CrossRef]
- Chen, Q.; Xiao, H. A neural knowledge graph evaluator: Combining structural and semantic evidence of knowledge graphs for predicting supportive knowledge in scientific QA. *Inf. Process. Manag.* 2020, 57, 102309–102315.
- Yang, J.B.; Liu, J.; Wang, J.; Sii, H.S.; Wang, H.W. A belief rule-base inference methodology using the evidential reasoning approach—RIMER. *IEEE Trans. Syst. Man Cybern. Part A Syst. Hum.* 2006, 36, 266–285. [CrossRef]
- Cao, Y.; Zhou, Z.J.; Hu, C.H.; He, W.; Tang, S.W. On the interpretability of belief rule based expert systems. *IEEE Trans. Fuzzy* Syst. 2021, 29, 3489–3503. [CrossRef]
- 29. Ikram, R.M.A.; Goliatt, L.; Kisi, O.; Trajkovic, S.; Shahid, S. Covariance Matrix Adaptation Evolution Strategy for Improving Machine Learning Approaches in Streamflow Prediction. *Mathematics* **2022**, *10*, 2971. [CrossRef]
- Kumar, P.; Kumar, P.; Hati, A.S.; Kim, H.S. Deep Transfer Learning Framework for Bearing Fault Detection in Motors. *Mathematics* 2022, 10, 4683. [CrossRef]
- Cheng, J.; Yan, X. Application of Teaching-Learning-Based Optimization Algorithm in Rotor Fault Diagnosis for Asynchronous Motor. Procedia Comput. Sci. 2018, 131, 1275–1281. [CrossRef]
- Al-Ameri, S.M.; Abdul-Malek, Z.; Salem, A.A.; Noorden, Z.A.; Alawady, A.A.; Yousof, M.F.M.; Mosaad, M.I.; Abu-Siada, A.; Thabit, H.A. Frequency Response Analysis for Three-Phase Star and Delta Induction Motors: Pattern Recognition and Fault Analysis Using Statistical Indicators. *Machines* 2023, 11, 106. [CrossRef]
- Kostal, T.; Kobrle, P. Induction Machine On-Line Parameter Identification for Resource-Constrained Microcontrollers Based on Steady-State Voltage Model. *Electronics* 2021, 10, 1981. [CrossRef]
- Chen, H.T.; Chai, Z.; Dogru, O.; Jiang, B.; Huang, B. Data-driven designs of fault detection systems via neural network-aided learning. *IEEE Trans. Neural Netw. Learn. Syst.* 2022, 33, 5694–5705. [CrossRef]
- Chen, H.; Luo, H.; Huang, B. Transfer Learning-motivated Intelligent Fault Diagnosis Designs: A Survey, Insights, and Perspectives. *TechRxiv* 2022. [CrossRef]
- Pei, J. A Dictionary-based Maximum Match Algorithm Via Statistical Information for Chinese Word Segmentation. Int. J. Electron. Inf. Eng. 2020, 12, 24–33.
- Li, P.; Guang, Y.; Qiao, T. Research on Chinese New Word Recognition Method. In Proceedings of the 2020 4th International Conference on Electronic Information Technology and Computer Engineering, Xiamen, China, 6–8 September 2020.

- 38. Jiang, S.; Fu, S.; Lin, N. Pretrained Models and Evaluation Data for the Khmer Language. *Tsinghua Sci. Technol.* **2022**, 27, 709–718. [CrossRef]
- 39. Che, J.L.; Tang, L.W.; Deng, S.J.; Su, X.J. Chinese Word Segmentation based on Bidirectional GRU-CRF Model. *Int. J. Perform. Eng.* **2018**, *12*, 3066. [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.