

Research Article

Evaluation Model of Power Operation and Maintenance Based on Text Emotion Analysis

Wei Wan ¹, Yuanlong Liu ^{1,2}, Xingwang Han ³ and Huijian Wang ³

¹Shandong SGIT Technology Co., Ltd., Jinan 250101, China

²Shandong University, Jinan 250104, China

³Information & Telecommunications Co., State Grid Shandong Electric Power Co, Jinan 250104, China

Correspondence should be addressed to Yuanlong Liu; liuyuanlong@sgitg.sgcc.com.cn

Received 15 November 2021; Accepted 13 December 2021; Published 29 December 2021

Academic Editor: Gengxin Sun

Copyright © 2021 Wei Wan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The application of data mining technology in power field mainly focuses on the application of power defect text and dispatching text. However, the power operation and maintenance data contains a lot of information about power equipment suppliers. Taking the operation and maintenance text involving power equipment suppliers as an example, this paper summarizes the theme of operation and maintenance text and studies the evaluation model of power equipment suppliers. The next sentence prediction analysis model of single round dialogue text based on transformer bidirectional encoder prediction and cosine similarity weighting is proposed, which can effectively divide the topic of dialogue text. Aiming at the semantic richness and complexity of power operation and maintenance text, a supplier evaluation model based on text emotion analysis is proposed. Based on the expansion of the entries and attributes of the existing power ontology dictionary, the dialogue emotion analysis rules are established to realize the normal evaluation of power equipment suppliers.

1. Introduction

At present, the basic requirement of ubiquitous power Internet of things construction is to realize business collaboration and data connectivity. With the improvement of the intelligence of power grid, the power field is facing the explosive growth of data information. Correctly processing and analyzing this data information and maximizing its value are among the biggest problems faced by power grid construction. A large proportion of unstructured data accumulated in the construction and development of smart grid, as an important part of power big data, is of great significance to the intelligent development of power grid.

In terms of the diversity of data structure types, big data in the power field mainly includes structured data and unstructured data [1]. Structured data refers to data with certain structure, which can be divided into fixed basic elements and can be represented by two-dimensional tables. Unstructured data is data other than structured data. The data structure is not fixed, including text, pictures, audio,

and video. Its characteristics are not easy to standardize, and it is difficult for computers to obtain its information directly. At present, there are mature ways to analyze and process structured data, but the method of mining unstructured data is still in its infancy. Intelligent mining of unstructured text data of power grid is a hot and difficult problem.

The intelligent mining of unstructured text data information in power field can use natural language processing technology [2]. The application of natural language processing technology in the field of electric power mostly focuses on text classification and knowledge Atlas, and there is little exploration in the direction of emotion analysis and information retrieval. At present, the application of text intelligent mining technology in power field mainly focuses on the application of power defect text and dispatching text, and there is little research on dialogue text.

In the power field, the management and technical personnel of power grid enterprises often mention the content of equipment suppliers in the process of work communication. For the evaluation needs of suppliers, the

dialogue text can be summarized by using text intelligent mining technology, and the dialogue text topics can be summarized by taking the supplier as the unit. The emotional analysis of the topic power dialogue text is carried out, the emotional tendency of the dialogue is judged, and the emotional score is obtained, the corresponding rules of emotional judgment and supplier evaluation are designed, and the power supplier evaluation model is established. Therefore, in this paper, the text intelligent mining in the power field is taken as the research object, the power text is analyzed and processed by using natural language processing technology, effective information is obtained, and support for the evaluation of power equipment suppliers is provided.

In recent years, the value of massive unstructured data in the power field has attracted more and more attention. Text intelligent mining provides a technical means for exploring and analyzing the potential value of text content. For the dispatching fault record, Zhang et al. [3] established the fault classification model of power grid inspection well by using database knowledge discovery and data mining. For equipment failure, replacement, maintenance, test, and other events and equipment characteristic text data, Anello and Del Rosso [4] used machine learning algorithm to realize fault prediction and preventive maintenance applied to power grid infrastructure. Regarding n order to solve the problem of complex modeling and difficult maintenance of power grid diagnosis system, Guo et al. [5] proposed a power grid remote signaling information analysis method based on semantic analysis. Wang and Xu [6] used k-nearest neighbor algorithm to classify the defect degree of power supply equipment text. Wang et al. [7] proposed a transformer state evaluation method based on multisource operation information to realize equipment health state detection. Liu and Singh [8] conducted in-depth research on power text data and proposed a research model of power reliability index based on power text mining and multisource data fusion. Chen et al. [9] proposed a defect text processing method based on recurrent neural network and long-term and short-term memory and realized the defect text classification of power equipment. Rocchetta et al. [10] studied the analysis models of classification, identification, retrieval, and evaluation of power grids defect operation and maintenance text. Zhang et al. [11] proposed a two-layer classification model based on machine learning to realize scheduling fault diagnosis based on alarm signal. Jin et al. [12] realized the automatic extraction of regulation text knowledge based on deep learning model in power grid regulation. Ban and Ning [13] established a manually labeled open-source power business intention identification data set for power business dialogue, which laid a foundation for the verification of power dialogue text classification model.

2. Topic Induction Method of Power Suppliers Based on Text Data Mining

In the process of power equipment supplier evaluation, power grid enterprises usually take performance, after-sales service, and quality supervision as the evaluation contents.

Some power grid companies require project units to evaluate suppliers' quality, service, and supply through manual statistics and direct scoring. However, this evaluation method is not inefficient and its accuracy is affected by the experience of evaluators and the scope of evaluation materials and also affected by the distortion of evaluation results caused by suppliers only paying attention to the evaluation content and neglecting other quality. Intelligent mining of power dialogue text can obtain the actual equipment operation quality and service level of suppliers from daily work practice.

According to the characteristics of dialogue text in power grids, in this paper, an evaluation method of power equipment suppliers considering the characteristics of power dialogue text is proposed. Based on the research of dialogue text features, a method of sentence prediction and cosine similarity weighting based on bidirectional encoder is proposed. The method judges the coherence of the upper and lower sentences of the power dialogue text and designs the supplier topic induction rules to realize the topic induction of the dialogue text through the research on the sentence matching of a single round of dialogue and the cross judgment of dialogue interruption.

2.1. Sentence Prediction Analysis of Single Round Dialogue Based on Text Mining. Single round dialogue text is a dialogue text with only two sentences from multiple rounds of dialogue text, which is characterized by obvious correlation. Firstly, a prediction model (PGPM) for predicting the coherence of upper and lower sentences is obtained through corpus training.

The model takes two dialogue texts as input; the first tag [CL] is added, and the tag [SP] is inserted as a special separator between the two conversation texts. Each item T_{tok} of input layer is the vector of each word in the input statement. N and M represent the number of words in two sentences, respectively, input $t_{1:N}$ represents each word sequence $t_1, t_2, \dots, t_{1:N}$ in the previous conversation, and $t'_{1:M}$ represents each word sequence $t'_1, t'_2, \dots, t'_{1:M}$ in the next conversation. The input layer of each pair of dialogue is added by marker word embedding, segment word embedding, and position word embedding to obtain one-to-one corresponding $E_{[CL]}, E_1, \dots, E_N, E_{[SP]}$ and $E'_1, \dots, E'_M, E_{[SP]}$.

The model output is shown as follows:

$$p = \text{softmax}(C \cdot W^T), \quad (1)$$

where p represents the prediction matching probability matrix of the next sentence, C is the final implicit value of the first flag [CL] of the model, and W^T represents the full connection layer weight matrix.

The model is actually a binary classification problem, so p is a two-dimensional vector, and the values represent the probability values predicted as 0 and 1 in the next sentence, that is, uncorrelated and correlated probabilities. The prediction probability P_{NS} of the next sentence takes the value representing the correlation of the two sentences in the vector.

The model is a depth feature prediction based on words, which virtually ignores the word level features. Cosine similarity [14] can explore the factors related to the two sentences from the word level. Therefore, the next sentence prediction algorithm based on PGPM model and cosine similarity weighting is proposed and constructed in this paper; the purpose is to integrate the advantages of depth distribution analysis and cosine similarity word level analysis, so as to improve the accuracy of next sentence prediction of dialogue text.

When cosine similarity method is used to judge the coherence of dialogue text, the repeated contents in the text are often used as the coherence evaluation standard. Therefore, for word segmentation of a single round of dialogue text, the cosine similarity of adjacent dialogues can be calculated by using the following formula:

$$S = \cos(U, V) = \frac{\sum_{i=1}^n u_i \times v_i}{\sqrt{\sum_{i=1}^n (u_i)^2} \sqrt{\sum_{i=1}^n (v_i)^2}} \quad (2)$$

where S represents cosine similarity of adjacent conversations, and $U = (u_1, u_2, \dots, u_n)$ and $V = (v_1, v_2, \dots, v_n)$ are the n -dimensional word frequency feature vectors obtained after the word frequency vectorization of two sentences. Its value is between -1 and 1; the smaller the similarity, the greater the distance.

Based on the depth features and language similarity features, this paper defines the semantic relevance matching degree of a single round of dialogue, which is defined as follows:

$$T = (1 - \vartheta)P_{NS} + \vartheta S, \quad (3)$$

where T is semantic relevance matching degree of single round dialogue, ϑ represents cosine similarity weight coefficient, and P_{NS} is the second classification value standard.

The value of T is a number greater than or equal to 0. The greater the value of T , the greater the matching correlation between the two sentences. If the value is greater than or equal to 0.5, the upper and lower sentences are considered relevant, the upper and lower sentences are classified as the same conversation topic, and if it is less than 0.5, it is considered irrelevant. Its significance is that it can integrate the depth feature and similarity feature, comprehensively consider the language relationship between upper and lower sentences, and improve the accuracy of matching judgment. The function of ϑ is to balance the weight ratio of depth feature and similarity feature, and the optimal model of single round dialogue text judgment can be obtained by optimizing the coefficient.

So far, the prediction analysis of the next sentence of a single round of dialogue is completed. Through this step, the correlation judgment between the texts of each single round of dialogue can be obtained. It can be considered that the two dialogue texts judged as relevant are the same topic, which can be summarized into the same topic, and the two dialogue texts judged as irrelevant are not the same topic. However, only considering the relevance of a single round of dialogue will lead to one-dimensional thematic induction results. It is

also necessary to discuss the topic continuation and integration caused by the topic interruption and intersection of multiple rounds of dialogue texts.

In the power dialogue text, the topic interruption and intersection of multiple rounds of dialogue text often occur. In order to realize the dialogue topic induction based on the sentence matching of a single round of dialogue, the dialogue interruption cross processing flow is designed. The core idea of this process is to judge the sentence matching relevance of a single round of dialogue on the basis of judging whether the round interval conditions and user ID information meet the induction requirements of multiple rounds of topics. After the interval condition of multiple rounds of dialogue, the adjacent single round of dialogue text caused by the interruption or intersection of dialogue topics can be filtered based on the round information, which is judged to be irrelevant, but the same topic still exists within the allowable range of subsequent rounds. Using the user ID information condition, two sentences of dialogue texts separated regardless of distance can be included in the relevance judgment process. This condition is conducive to screening out the two sentences linked through the user ID to avoid being ignored due to the round interval. Finally, judge the matching relevance of the next sentence in a single round of dialogue. This step is the basic method of this process. The result is the basis for judging whether a single round of dialogue is the same topic.

2.2. Supplier Topic Induction Rules. After obtaining the same topic conversation set, according to the supplier information category in the power business ontology dictionary, the supplier information in each topic conversation set is extracted, the implicit evaluation object is identified by the upward proximity principle, and then the irrelevant redundant topic content is removed. The upward proximity principle means that when multiple supplier information appears, the topic of the dialogue text belongs to the previous supplier topic before the new supplier information appears.

The following three rules are adopted:

- (1) If the supplier information is not identified, it is considered that the conversation set is talking about irrelevant redundant content, which is worthless to the evaluation of equipment suppliers and can be screened out.
- (2) If one supplier information or multiple identical supplier information is identified, it is considered that the evaluation object of the dialogue set is the identified supplier.
- (3) There are two or more of the different supplier information, such as supplier A, B, \dots, X in the order of occurrence. For each text in the set, determine the corresponding supplier based on the principle of upward proximity, and define that the evaluation object of the dialogue set is supplier A from the first sentence to supplier B , the evaluation object is supplier B before the sentence from supplier B to

supplier C, and so on. If a supplier appears repeatedly, the dialogue sets of the supplier are merged.

So far, the topic text induction of each supplier is realized. The following is a specific example to illustrate the specific topic text induction process. The data set selects 361 pairs of single round conversation texts from the group conversation texts of power collection, operation and maintenance, and the cosine similarity weight coefficient ϑ in the semantic relevance matching degree T of a single round of dialogue is optimized; the accuracy is shown in Figure 1.

As shown in Figure 1, when ϑ is taken as 0.05, the maximum judgment accuracy of single round dialogue is 82.6%, and when it exceeds 0.05, the accuracy decreases monotonically.

LSTM + softmax model turns the output of neural network into a probability distribution; that is, each output is a decimal between (0 ~ 1), and the sum of all output results is 1. Through such processing, we can easily use the cross entropy loss function to calculate the distance between the real distribution and the expected distribution and use the gradient descent algorithm to reduce the cross entropy loss to fit the sample training model. Then, the classification model of LSTM followed by softmax [15] is used to compare the classification effects of the same topic in the upper and lower sentences, and the model accuracy index is shown in Table 1.

As shown in Table 1, the PGPM model in this paper can improve the accuracy of single round dialogue judgment. The establishment of the weighted model also has certain interpretability. The same text content in the dialogue process is more likely to be the discussion of the same topic.

Based on the prediction analysis of the next sentence of a single round of dialogue text, the cross interruption of dialogue is verified by experiments. The dialogue of the experimental data set includes two dialogue topics, namely, the discussion topic of abnormal meter counting supplier and the discussion topic of meter box cost performance; that is, when the staff discuss the topic of abnormal meter counting supplier, an irrelevant discussion topic of meter box price performance is inserted, so as to clearly extract the content related to supplier evaluation; the above two topics need to be identified and divided, and the dialogue text related to the meter box irrelevant to the supplier evaluation needs to be removed. The division of dialogue cross interrupt processing under various models is shown in Table 2.

It can be seen from Table 2 that, in terms of multiple round dialogue topic division, the model using cosine similarity calculation only has the lowest accuracy, and the weighted model (PGPM) has the highest accuracy. The PGPM model can correctly identify the corresponding supplier information under the same topic through the supplier identification rules and delete the redundant dialogue content from the dialogue topic.

3. Supplier Evaluation Model Based on Text Emotion Analysis

In the power field, it has been applied to mining equipment defects, operation and maintenance, and alarm signal text, which reduces the workload of manual defect analysis and improves the accuracy of analysis to a certain extent. Compared with the text that objectively describes power equipment defects, operation and maintenance, and alarm signals, words with emotional tendency will appear in the dialogue. Therefore, the mining of power dialogue text needs to be combined with emotional analysis.

Emotion analysis, also known as comment mining or opinion mining [16, 17], can be divided into affective dictionary based methods and machine learning based methods. There are various syntactic forms of power dialogue text, and the semantic emotion of power dialogue text is rich, and emotional transformation exists. According to the characteristics of dialogue text, firstly, the expansion of power ontology dictionary is studied. Because the power dialogue text is highly professional and has the characteristics of common abbreviations, it is different from the common thesaurus. In order to improve the accuracy of text understanding, it is necessary to establish an ontology dictionary in the field of dialogue business.

This paper adopts the semisupervised method, based on the domain ontology dictionary and general dictionary, takes the dialogue power business domain text as the corpus, carries out word segmentation based on Hidden Markov model (HMM) [18], selects new domain ontology words according to word frequency, and manually checks whether they become ontology words and synonyms to supplement the ontology dictionary. The new ontology dictionary categories in this paper mainly include supplier name, business domain events, and other words and expand supplier name and its common abbreviations to deal with the prominent characteristics of colloquial dialogue text.

Emotion analysis can usually be divided into two ideas: one is rule-based method, and the other is machine learning method. The rule-based method mainly compares the emotional words, emotional polarity, and other information in the text on the basis of the emotional dictionary after preprocessing the text, so as to calculate the emotional score in the text and judge the emotional tendency and intensity. The method based on machine learning is to extract the features of the text, quantify it, and select the classification algorithm to judge the emotional tendency of the text. In this paper, the power supplier evaluation model based on text emotion analysis technology is established based on rule-based method. It is mainly because the evaluation rules are formulated on the basis of a large number of studies on the characteristics of power dialogue content, and relatively reasonable parameter values can be set.

In the emotional analysis of power dialogue text, considering the basic information such as emotional keywords and word polarity, the cumulative method is used to judge the emotional tendency of dialogue text. The most important

FIGURE 1: Influence of θ on model accuracy.

TABLE 1: The accuracy of single-round dialogue text judgment.

Model	Accuracy of model (%)
Cosine similarity model	57.56
LSTM + softmax	76.35
PGPM model	80.67

step of emotion analysis of power dialogue text is to find the key words related to emotion in the sentence.

Through the observation and analysis of a large number of power dialogue texts, it is found that the main words affecting the emotional evaluation in the text are event word I_{ec} , emotion word E_{cn} , conjunction C_{nw} , degree adverb A_{dc} and negative word N_{uc} . These words have certain rules to follow for the function of the evaluation object. When the power domain ontology dictionary is expanded, the event word I_{ec} and emotion word E_{cn} have been labeled and set as the basic unit of emotion evaluation. At the same time, conjunction C_{nw} , degree adverb A_{dc} and negative word N_{uc} are selected as auxiliary units to modify the text emotion analysis.

For the subject text of specific suppliers, word segmentation, part of speech retrieval, and labeling are carried out for each sentence to obtain keyword information, and redundant sentences are deleted, and emotional scoring is carried out according to the following rules:

- (1) In a single sentence dialogue text, when at least one of the basic unit event words I_{ec} and emotional words E_{cn} appears, the sentence is considered to have emotional tendency and evaluation opinions; otherwise, the sentence is considered neutral.
- (2) When there is only event word I_{ec} in the sentence, it shall be included in the event score. When there is only emotional word E_{cn} in the sentence, it is included in the emotional score. When event word I_{ec} and emotional word E_{cn} coexist in a sentence, they are included in event score and emotion score respectively.

- (3) Event word I_{ec} is highly professional and mostly describes the occurrence of fault. I_{eci} is the emotional score of the i -th event word.
- (4) The emotional word E_{cn} refers to the emotional dictionary and sets the score according to the positive and negative emotions. E_{cnj} is the emotional score of the j -th emotional word.
- (5) Degree adverb A_{dc} sets the polarity coefficient according to the polarity intensity, taking 0.5, 1 or 2.
- (6) The negative coefficient of negative word N_{uc} is -1, and N_{ucl} is the l -th negative coefficient of the first negative word. If there is no negative word, $N_{uc0} = 1$.
- (7) When there are multiple event words I_{ec} or emotional words E_{cn} in a sentence, it is necessary to judge the emotional semantics and transformation of the sentence. When there is conjunction C_{nw} , the emotional semantic evaluation priority is determined according to the positional relationship between the conjunction and the basic unit and the conjunction turning relationship. If there is no conjunction in the sentence, $\beta_C = 1$ for all basic units.
- (8) When there are both event words I_{ec} and emotion words E_{cn} in a sentence, event words, as words that objectively state facts, have higher evaluation persuasion than subjective emotion words. The priority coefficient β_Y of basic unit relationship evaluation of event word I_{ec} and emotion word E_{cn} is 1.2 and 0.8, respectively.
- (9) For each basic unit in the sentence, combining the above rules (7) and (8), the evaluation priority coefficient is defined as

$$\beta = \beta_C \times \beta_Y \quad (4)$$

- (10) Regarding the attachment relationship between basic unit event word I_{ec} and emotional word E_{cn}

TABLE 2: The topic division of different models in multiple rounds dialogue text.

Model	Topic division of multiple round dialogue		Accuracy (%)
	Dialogue set	Supplier	
None	(5)	A	
	(8)	B	
	(9)	A	
	(1)/(2)/(3)/(4)/(6)/(7)/(10)/(11)	Redundancy	
Cosine similarity model	(5)/(6)/(7)/(8)/(9)/(10)/(11)	A, B	52.17
	(1)/(2)/(3)/(4)	Redundancy	
LSTM + softmax	(5)/(6)/(7)	A	82.75
	(8)/(9)/(10)/(11)	B	
	(1)/(2)/(3)/(4)	Redundancy	
PGPM model	(5)/(9)	A	100
	(8)	B	
	(1)/(2)/(3)/(4)/(6)/(7)/(10)/(11)	Redundancy	

and degree adverb A_{dc} and negative word N_{uc} , if there is conjunction C_{mw} , it is divided by conjunction; if there is no conjunction, it is divided by each basic unit for emotional judgment.

For the same topic dialogue set, the supplier score shall be given in combination with the relationship characteristics of the basic unit and auxiliary unit of emotional evaluation. Each basic unit is directly affected by the priority coefficient, degree adverb, and negative word coefficient of the unit, so

the score of the basic unit is directly corrected by the coefficient series product. For suppliers with the same subject and the same evaluation object, their emotion score value in sentences is calculated. The positive and negative of the value can directly reflect the positive or negative evaluation of the interlocutor, and the size of the value reflects the intensity of emotion. Thus, a single sentence supplier score as shown in the following formula is established:

$$V_{SN_u} = \sum_{i=1}^I \left(\beta_{Ii} \times \prod_{k=0}^K A_{dk} \times \prod_{l=0}^L N_{ucl} \right) + \sum_{j=1}^J \left(\beta_{Sj} \times \prod_{t=0}^T A_{dt} \times \prod_{q=0}^Q N_{ucq} \right), \quad (5)$$

where V_{SN_u} represents single sentence supplier score, I represents the number of event words in the sentence, K represents the number of adverbs of degree attached to each event word, L represents the number of subsidiary negatives of each event word, J represents the number of emotional words in a sentence, T represents the number of degree adverbs attached to each emotional word, and Q represents the number of negative words attached to each emotional word.

The supplier's overall evaluation is divided into the sum of all single sentence scores:

$$V_{SN} = \sum_{u=1}^U V_{SN_u}, \quad (6)$$

where V_{SN} represents total supplier evaluation score, and U is the number of supplier evaluation sentences.

This paper takes the dialogue text in the field of power acquisition and operation and maintenance as an example to verify the proposed model. The ontology words are marked with entry attributes, synonyms, and synonyms. The attributes include proprietary domain nouns, supplier names, and event key words. In order to show the emotional analysis process of dialogue text, the sentence "Supplier A used to be very good, but now why does the clock often come out

abnormal?" is taken as an example for keyword extraction. On the basis of ontology dictionary and general dictionary in the field of power collection, operation and maintenance, text preprocessing, word segmentation, and keyword extraction and annotation are carried out for the text. The key words of emotional analysis of dialogue text are shown in Table 3.

On this basis, a horizontal comparison is made. Taking semantic transformation, semantic extension, and ordinary single sentences as examples, the judgment accuracy of the methods in this chapter for semantic transformation is tested, and the analysis and scoring results are shown in Table 4.

Using the subject division results, different models are used to analyze the data set to verify the supplier evaluation effect. The results are shown in Table 5. Using the subject division results, different models are used to analyze the data set to verify the supplier evaluation effect. The results are shown in Table 5.

It can be seen from Table 5 that the scoring results obtained by the proposed model and LSTM + softmax model are the same due to the consistent subject division, and it is quite different from those models under the nonprocessing and cosine similarity calculation models. For scoring of supplier A, the nonprocessing model only selects the

TABLE 3: The key words of emotional analysis of dialogue text.

Supplier	I_{ec}	E_{cn}	C_{mw}	A_{dc}	N_{uc}
A		Good	But	Very	
A	Clock			Often	

TABLE 4: Emotional analysis and semantic comparison score of dialogue text.

Sentence	Score
It used to be very good, but now why do you often come out with abnormal clock?	-2.15
It used to be good, and now it's good.	1.25
It used to be good.	0.75

TABLE 5: Dialogue text emotion analysis supplier score.

Model	Score	
	Supplier A	Supplier B
None	0.85	-0.75
Cosine similarity model	0.15	-0.75
LSTM + softmax	-2.85	-0.75
PGPM model	-2.85	-0.75

sentences containing information of supplier A as the scoring object, and other relevant sentences are missing, while the cosine similarity model also lacks some information, resulting in poor scoring results. The score of supplier B was not affected, because only one conversation was involved. It can be seen that the supplier evaluation results based on emotion analysis are greatly affected by the correctness of subject division. The accuracy of weighted model in subject division can effectively improve the authenticity and reliability of supplier evaluation.

4. Conclusions

A large proportion of unstructured data accumulated in the development of smart grid. As an important part of power big data, data mining of unstructured text data of power grid is a hot issue today. This paper takes the application of text intelligent mining technology in power field as the research object, analyzes the dialogue text involving power equipment suppliers, and establishes the topic induction model of power dialogue text and the evaluation model of power suppliers. The experimental analysis shows that the evaluation of power equipment suppliers based on dialogue text intelligent mining is feasible and effective and can be used as a useful supplement to the current evaluation methods. At present, there are few researches on unstructured and structured data fusion. It is very important to broaden the research perspective of unstructured data mining. How to break the traditional thinking of data analysis and carry out extensive research combined with text intelligent mining is an important direction of future research.

Data Availability

The basic data used in this paper were downloaded from the online public data set: ARPA-E GRID DATA (<https://db.bettergrids.org/bettergrids/community-list>).

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

- [1] N. Vo, S. Liu, and X. Li, "Leveraging unstructured call log data for customer churn prediction," *Knowledge-Based Systems*, vol. 212, no. 4, Article ID 106586, 2021.
- [2] R. Lou, D. Lalevic, C. Chambers, H. M. Zafar, and T. S. Cook, "Automated detection of radiology reports that require follow-up imaging using natural language processing feature engineering and machine learning classification," *Journal of Digital Imaging*, vol. 33, no. 1, pp. 131–136, 2020.
- [3] T. Zhang, Z. Chen, Z. Zhao, X. Luo, W. Zheng, and W. Chen, "FaultTracer: interactive visual exploration of fault propagation patterns in power grid simulation data," *Journal of Visualization*, vol. 24, no. 5, pp. 1051–1064, 2021.
- [4] M. T. Anello and A. Del Rosso, "Probabilistic assessment of wind farm disconnection by power grid failures," *IEEE Latin America Transactions*, vol. 14, no. 4, pp. 1808–1815, 2016.
- [5] S. Guo, H. Zhao, and H. Zhao, "The most economical mode of power supply for remote and less developed areas in China: power grid extension or micro-grid?" *Sustainability*, vol. 9, no. 6, 910 pages, 2017.

- [6] W. Wang and X. Xu, "Power electronic signaling technology—a new class of power electronics applications," *IEEE Transactions on Smart Grid*, vol. 3, no. 3, pp. 332–339, 2010.
- [7] H. Wang, G. Ni, and J. Chen, "Research on rolling bearing state health monitoring and life prediction based on pca and internet of things with multi-sensor," *Measurement*, vol. 157, no. 4, Article ID 107657, 2020.
- [8] Y. Liu and C. Singh, "A methodology for evaluation of hurricane impact on composite power system reliability," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 145–152, 2011.
- [9] K. Chen, R. J. Mahfoud, and Y. Sun, "Defect texts mining of secondary device in smart substation with glove and attention-based bidirectional lstm," *Energies*, vol. 13, no. 17, pp. 1–17, 2020.
- [10] R. Rocchetta, L. Bellani, M. Compare, E. Patelli, and E. Zio, "A reinforcement learning framework for optimal operation and maintenance of power grids," *Applied Energy*, vol. 241, pp. 291–301, 2019.
- [11] X. Zhang, S. Yue, and X. Zha, "Method of power grid fault diagnosis using intuitionistic fuzzy petri nets," *IET Generation, Transmission & Distribution*, vol. 12, no. 2, pp. 295–302, 2018.
- [12] W. C. Jin, S. Y. Heo, and M. K. Kim, "Hybrid operation strategy of wind energy storage system for power grid frequency regulation," *IET Generation, Transmission & Distribution*, vol. 10, no. 3, pp. 736–749, 2016.
- [13] H. Ban and J. Ning, "Design of English automatic translation system based on machine intelligent translation and secure internet of things," *Mobile Information Systems*, vol. 2021, Article ID 8670739, 8 pages, 2021.
- [14] G. Sun and C. C. Chen, "Influence maximization algorithm based on reverse reachable set," *Mathematical Problems in Engineering*, vol. 2021, pp. 1–12, Article ID 5535843, 2021.
- [15] A. S. Devi, G. Maragatham, K. Boopathi, and A. G. Rangaraj, "Hourly day-ahead wind power forecasting with the EEMD-CSO-LSTM-EFG deep learning technique," *Soft Computing*, vol. 24, no. 16, pp. 12391–12411, 2020.
- [16] E. M. L. Barrón, C. R. Zatarain, B. R. Oramas, and M. Graff, "Opinion mining and emotion recognition applied to learning environments," *Expert Systems with Applications*, vol. 150, Article ID 113265, 2020.
- [17] J. B. Awotunde and R. G. Jimoh, "A model for identifying influential bloggers using social proof, mining comment, and emerging topics in social media," *Computing and Information Systems*, vol. 23, no. 1, pp. 19–25, 2019.
- [18] K. Jahromi, D. Gharavian, and H. Mahdiani, "A novel method for day-ahead solar power prediction based on hidden Markov model and cosine similarity," *Soft Computing*, vol. 24, no. 7, pp. 4991–5004, 2020.