

Transformer fault text classification model based on double feature channel

Fujun Guan^{1,*}, Dewen Wang^{1,2}

¹ College of Control and Computer Engineering North China Electric Power University, Baoding 071003, China

² Engineering Research Center of Intelligent Computing for Complex Energy Systems, Ministry of Education, Baoding 071003, China

* Corresponding author: Fujun Guan

Abstract: The power transformer plays a key role in the power distribution and transmission of the power system of the equipment, if the transformer fault leads to unexpected outage of the power system, then it will have a great threat to the whole power system, through the analysis of the transformer fault text can better guide the procurement of the transformer and fault maintenance. This paper proposes a classification method of transformer fault text based on two feature channels. This method extracts useful text emotion information from different aspects and levels by using two feature vectors, namely word vector and pinyin vector, and inputs it into Bilstm network to extract local and global semantic features. Then the features are fused through the attention mechanism. Finally, the features output by softmax function are classified and predicted. Experimental results show that the proposed method greatly improves the accuracy of transformer fault text classification.

Keywords: Transformer; Fault text; Attention mechanism; Double feature channel.

1. Introduction

Power transformer is an important equipment in the power grid, supporting the safe operation of the power system, no matter the transformer in the operation of any problems, will bring great economic property losses to the national home, and even endanger people's personal safety, have a serious social impact. Therefore, it is of great significance to use the existing transformer defect text to accurately and quickly classify the transformer fault category to guide the transformer procurement, installation, daily maintenance, rapid fault location and timely maintenance [1-2].

The characteristics of power transformer fault text are very different from the general text. Firstly, there are a large number of professional words in power transformer fault text. Generally, the transformer fault text recorded by power inspection personnel is short text, which is different from the long text. In addition to the professional words, there are also some professional symbols or letters in the text to describe the failure of transformer equipment. Therefore, there are some difficulties in mining power transformer fault text.

2. Research Status

Early researchers mostly integrated structured data information to carry out state assessment [2-8]. However, transformer fault texts were generally recorded by transformer managers in the daily maintenance process, in the form of Chinese handwritten or electronic reports, transformer fault phenomena, fault causes, operation, maintenance and overhaul methods and other unstructured data. Therefore, the early methods could not accurately identify these unstructured data. At present, machine learning, natural language processing and other technologies are gradually applied to the task of transformer fault text classification. Jiang Yiwen et al. [9] combined the different characteristics of CNN and RNN to establish a circular convolutional neural network (RCNN) network, and added a maximum pooling layer to RCNN to extract the characteristic

information of transformer defect text. Although some classification results have been achieved, further improvement is needed because it is impossible to learn the hidden semantic information from other perspectives of the defective text.

At present, the domestic research on text data mining in the field of power is still in the preliminary exploration stage. Chinese text paragraph and English text paragraph have great differences in grammar structure, language habits, parts of speech and so on, English text mining method can not be directly applied to Chinese text. In addition, Chinese power text also contains equipment number, abbreviation and many other professional words, so it will be more difficult to carry out data mining. Literature [10] constructs the lifetime state evaluation model of circuit breakers through text mining, and realizes the similarity study of circuit breaker defects by mining the text of power grid defects. Literature [11] diagnoses power dispatching faults by mining alarm signal texts. Literature [12] builds a marketing management decision tree mining model according to the characteristics of existing power grid marketing data, so as to deduce accurate power grid marketing behaviors, with better prediction accuracy and lower data consumption. Literature [13-15] studies the power grid fault diagnosis method based on big data analysis, synthesizes a variety of information sources such as power grid topology information, circuit breaker trip information and relay protection device action information, and considers the loss of information to realize effective fault diagnosis.

3. Methods

Due to the inherent characteristics of transformer fault text data, the semantic features of transformer fault text data mining model for power industry are not comprehensive enough. Unable to follow key information in the fault text. Unable to obtain enough deep semantic information in the grid equipment fault text. And because the power grid equipment fault text is generally recorded by the maintenance personnel, there are mistakes, pinyin, homophone and other defects. Some models cannot accurately identify error

information such as misspellings. These errors and differences can not be identified effectively, which will affect the accuracy of the model to identify fault types. Based on this, this paper proposes a text recognition model for transformer faults based on double feature channels. This model can effectively solve the context-dependent problem and extract the deep feature information of the fault text. At the same time, it can also solve the problem that the previous model cannot accurately identify wrong characters by pinyin vector.

3.1. Dual-feature channel

Firstly, the word vector and pinyin vector of the collected transformer fault text should be extracted.

Word vector uses word2vec model to generate word vector. word2vec word vector can better consider contextual semantic information and avoid dimension "disaster" problem. The main text representation methods include thermal coding, TF-IDF and so on. The advantages of thermal coding are simple in principle and convenient in use. The disadvantage is that the semantically relevant information is not taken into account, which easily leads to the problem of dimensional "disaster". The use of word2vec word vector can be a good solution to this problem.

Pinyin feature vector is to convert the original sentences and symbols of transformer fault text using Pinyin. And the converted pinyin is represented vectorically by a large corpus. By using pinyin feature vector, the representation of the text is expanded, the feature information of the text is further enriched, and the influence of the wrong characters and homophones in the transformer fault text on the fault type identification results is reduced.

3.2. Bilitm model

Long - short - term memory neural network is based on cyclic neural network improvement. When processing long sequences of text, RNN can only obtain the information of relatively recent sequences due to the existence of gradient dispersion, and does not have the memory function of early sequences, so a lot of information is lost. By adding cell states and threshold mechanisms, LSTM solves the problems of gradient dispersion, gradient explosion, and only having short-term memory of RNN. LSTM consists of oblivion gate, input gate and output gate. The forgetting gate determines whether the information needs to be retained through the sigmoid function.

BiLSTM is a variant of LSTM, which combines forward propagation LSTM and back propagation LSTM to obtain semantic features in both directions. Bilstm can better capture bidirectional semantic dependence and solve the problem of context dependence. The bidirectional cyclic neural network model structure is shown in Figure 1 below.

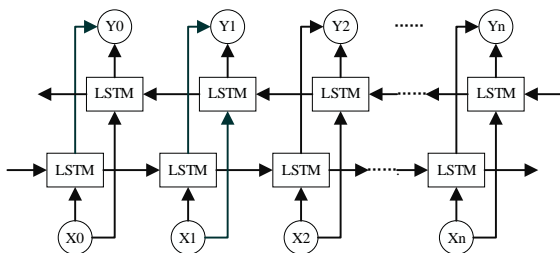


Figure 1. Schematic diagram of bidirectional cyclic neural network model

3.3. Attention mechanism

Attention mechanism is a method to extract important features of data, which was initially applied to image processing tasks. Bahdanau et al. first proposed the application of attention mechanisms in natural language processing tasks to alleviate the problem of long-distance dependence of RNNs. In text analysis, attention mechanism can solve the problem that source sequence and target sequence cannot establish dependence because of the long distance. Due to the different levels of importance of words and the different weight of features in the text, the introduction of attention mechanism can better learn the dependency relationship between words and enhance the attention to important words. The attention mechanism model is shown in Figure 2 below.

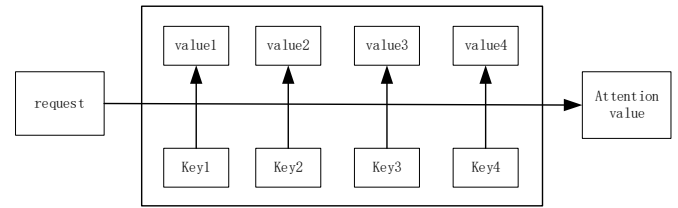


Figure 2. Attention mechanism model

The use of attention mechanism can help to increase the weight of key information of transformer fault text classification and reduce the weight of interference nonsense information, so as to further improve the accuracy of model feature extraction. The attention mechanism carries out weighted transformation on the state information sequence (hidden layer weight) extracted by BiLSTM, and automatically assigns different weights according to the importance of information.

3.4. Model of this paper

The model diagram of the model in this paper is shown in Figure 3 below.

Firstly, the collected text data of the main transformer fault is preprocessed. In this paper, the stuttering word segmentation tool is used for word segmentation. In the process of word segmentation, punctuation and stop words in the text are filtered.

The processed data uses the word2vec model to generate word vectors. At the same time, these fault texts are converted using Chinese pinyin. And the converted pinyin is represented vectorically by a large corpus.

Then, the generated word2vec word vector features and pinyin features are used as two different feature channel vectors, and then these vectors are input to BiLSTM layer to extract the local and global semantic features of transformer fault text information.

The self-attention mechanism is introduced into the model to fully extract the important semantic features in the text.

Then, the features extracted by the attention mechanism layer are spliced and fused, and the output features of the fusion layer are classified and predicted by the softmax function.

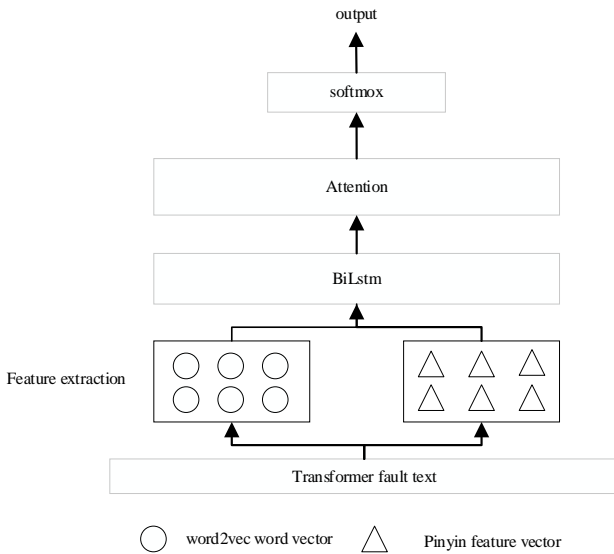


Figure 3. Text model diagram of transformer fault with double feature channel

4. Experiment

4.1. Data Set

780 transformer failure texts from all over the country were collected. According to the fault standards, the transformer fault text information is divided into three types: general, major, and emergency according to different fault degrees. The processed data set is divided into training set, test set and verification set according to the ratio of 7:2:1.

4.2. Algorithm evaluation index

In this paper, Precision, Recall and F1 evaluation indexes were respectively used to evaluate the model performance. Where, let A, B, C and D represent the number of positive, false negative, false positive and positive negative categories respectively.

The formula of precision rate P is:

$$P = \frac{A}{A+B} \quad (1)$$

The formula of recall rate R is:

$$R = \frac{A}{A+C} \quad (2)$$

Formula 1 value can be calculated as follows:

$$F1 = \frac{2 \cdot R \cdot P}{R+P} \quad (3)$$

4.3. Comparative experiment

In order to verify the performance of MFFM, the three classical models in this paper and short text classification, SVM, BILSTM and TextCNN are used as comparison models for comparison experiment, and the experimental results are shown in Table 1.

Table 1. Experimental results

Model	P	R	F1
SVM	0.6377	0.6256	0.6316
BILSTM	0.7476	0.7269	0.7371
TextCNN	0.7628	0.7403	0.7513
Model of this paper	0.8426	0.8389	0.8407

Through the analysis of Table 1, it can be seen that the classification performance of the transformer fault text data set of the proposed model is superior to the other three models.

As can be seen from the above table, F1 value of SVM model is only 0.6316. BILSTM

The F1 value is 0.7371. The F1 value of TextCNN is 0.7513, slightly higher than that of BILSTM. However, F1 value of this model is 0.8084, 8.94% higher than that of TextCNN model. In addition, from the perspective of accuracy and recall rate, the model in this paper is superior to other models.

5. Conclusion

Aiming at the problem that the existing transformer fault type identification methods are not comprehensive enough to obtain the semantic features of the fault text and cannot make full use of the deep semantic information of the fault text. In this paper, a text type recognition model of transformer based on double feature channel is proposed.

Although the recognition effect of the proposed model is better than that of other reference models, it still has some shortcomings, such as not adding dependency syntax text information features to identify transformer fault types. Further work will consider adding dependency syntax vectors and so on.

References

- [1] Hu Yizhuo, Xie Jiacheng, Li Jinzhong et al. Optimization of transformer winding deformation criterion considering insulation aging or damp [J]. Power grid technology, 2020, 44(06): 2361-2371
- [2] Hu Lijuan, Diao Yinglong, Liu Xscientific et al. Reliability analysis of distribution network operation based on big data technology [J]. Power grid technology, 2017, 41(01): 265-271
- [3] Liu Kezhen, Gou Jiutuo, Luo Zhao, et al. Prediction method of dissolved gas concentration in transformer Oil based on Particle Swarm Optimization and long short-term Memory Network model [J]. Power grid technology, 2020, 44(07): 2778-2785.
- [4] Zhang Peng, Qi Bo, Zhang Ruoyu et al. Prediction Method of Dissolved Gas in Transformer Oil Based on Empirical Wavelet Transform and Gradient Lifting radial basis [J]. Power grid technology, 2021, 45(09): 3745-3754
- [5] Geng Qishen, Wang Fenghua, Jin Xiao. Sound Diagnosis of Mechanical Fault of Dry Transformer Based on Gammatone Filter cepstral Coefficient and Whale Algorithm Optimized Random Forest [J]. Electric power automation equipment, 2020, 40(08): 191-196
- [6] Liu Yunpeng, Wang Bowen, Yue Haotian, Gao Fei, Han Shuai, Luo Shihao, Zhang Chenchen. Transformer biased magnetic Voseprint Recognition Based on 50Hz cepstral coefficient and gated cyclic Unit [J]. Proceedings of the csee, 2020, 40(14): 4681-4694.
- [7] Xianhua He, Yuanpeng Zhang, Guixing Cui, Bo Li, Xiaoyan Liang, Han Wang. Research on transformer fault detection method based on regression algorithm [J]. Protection and control of power systems, 2020, 48(21): 132-139.
- [8] Zhang L J, Sheng G X, Hou H J, et al. Interturn Short Circuit Fault Identification Method of Oil-Immersed Transformer Based on Fusion Analysis of Electric Heating Characteristics [J]. Power grid technology, 2021, 45(07): 2473-2482.
- [9] Jiang Yiwen, LI Li, LI Zhiwei, et al. Power Transformer Operation Text Information Mining Method Based on Deep Semantic Learning [J]. Proceedings of the csee, 2019, 39 (14): 4162-4171.
- [10] Qiu Jian, Wang Huifang, Ying Gaoliang, Zhang Bo, Zou Guoping, He Pentium. Text Information Mining Technology and its Application in Circuit Breaker Lifetime Condition

- Evaluation [J]. Automation of Electric Power Systems, 2016, 40(06): 107-112+118.
- [11] Wang Cuiyang, Jiang Quanyuan, Tang Yajie, Zhu Bingquan, Xiang Zhongming, Tang Jian. Based on text mining alarm signal fault diagnosis [J]. Electric power dispatching of electric power automation equipment, 2019, 33 (4) 6:126-132. The DOI: 10.16081 /j.i SSN. 1006-6047.2019.04.019.
- [12] LAI Guang-yuan, Li Jialiang. Grid Business marketing Data Mining Method based on Feature awareness iteration [J]. Computer Applications and Software, 2017, 34(08): 76-80.
- [13] ZHANG Chengjun, YIN Yan, BAO Jiusheng, et al. Research progress in fault diagnosis methods based on multi-source information fusion[J]. Journal of Hebei University of Science and Technology, 2014, 35(3): 213-221.
- [14] ZHANG Ming, JIANG Zhinong. Reciprocating compressor fault diagnosis technology based on multi-source information fusion[J]. Chinese Journal of Mechanical Engineering, 2017, 53(23): 46-52.
- [15] Zhao Wenqing, Zhu Lingyu, Gao Shuguo, Li Gang. Power transformer fault diagnosis based on multi-source information fusion method research [J]. Electric power information and communication technology, 2018 (10): 25-30. DOI: 10.16543/2095-641 - x. J. Electric power. Ict. 2018.10.005.