

Received July 19, 2020, accepted September 2, 2020, date of publication September 14, 2020,
date of current version September 24, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3024060

An Automatic Mapping Method of Intelligent Recorder Configuration Datasets Based on Chinese Semantic Deep Learning

JIANGBO REN¹, TIECHENG LI², SHAOBO GEN¹, QINQUAN LIU², YAKUN HE¹,
ZHIHUA WANG³, AND DAMING ZHOU³

¹State Grid Hebei Electrical Power Company, Shijiazhuang 050021, China

²State Grid Hebei Electrical Power Company Research Institute, Shijiazhuang 050021, China

³Wuhan Kemov Electrical Company Ltd., Wuhan 430023, China

Corresponding author: Daming Zhou (zdm2941894638@sina.com)

This work was supported by the Science and Technology Program of Chinese State Grid Corporation under Grant 5100-202013020A-0-0-00.

ABSTRACT Mapping the Intelligent Electronic Devices (IEDs) output interface address description datasets to the intelligent recorder is the groundwork for the recorder to accurately collect IEDs' operation information. These datasets, which are also intelligent recorder configuration datasets, are included in the Substation Configuration Description (SCD) file of an intelligent substation. The mainstream mapping method is manually mapping these datasets based on output interface Chinese description texts. When the number of IEDs is extremely large, the manual operation often takes a huge amount of time, together with the higher labor cost. And since the Chinese description texts have a certain degree of irregularity, it also poses a problem for the automatic mapping of the datasets. Aiming at this problem, this paper proposes an automatic mapping method of IEDs configuration datasets based on a deep learning framework—Dynamic Convolutional Neural Network (DCNN). Firstly, it uses the word representation model Word2vec to vectorize words in Chinese description texts as well as their semantics relationships. Then word vectors will be imported in the DCNN, which, based on its multilayer abstract learning characteristics of typical sample features, can perform semantic law mining and automatic mapping. The configuration datasets of intelligent recorder will be automatically mapped based on the Chinese descriptions mapping result. The Practical example shows that the Chinese description texts classification method based on the Dynamic Convolutional Neural Network model has strong semantic analysis ability and high classification accuracy, which effectively improves the accuracy of automatic mapping of intelligent recorder configuration data.

INDEX TERMS Intelligent substation, substation configuration description, intelligent electronic device, dynamic convolutional neural network.

I. INTRODUCTION

The intelligent recorder, which integrates the functions of transient signal recording, network message recording, on-line monitoring, diagnosis of secondary equipment, and managing the fault information for relay protection, is an important device for the operation and maintenance of the intelligent substation. The substation configuration description (SCD) file which is written according to the IEC61850 protocol is the key input file for an intelligent recorder to accurately monitor Intelligent Electronic Devices

(IEDs) operation information. When the recorder is put into operation, it collects the IEDs operation information and divides them into three different information groups for classified monitoring, including the strap information group, alarming information group and status monitoring information group. Each one contains sub information groups, such as hard strap information group, SV receiving strap information group and function strap information group, etc. Mapping the information of each IED data output interface address in SCD file to different information groups of the recorder is the basic step to ensure that the operation recorder can monitor different types of IED operation information accurately in real time. Therefore, the IED data output interface address

The associate editor coordinating the review of this manuscript and approving it for publication was Shuaihu Li¹.

description datasets are also the intelligent recorder configuration datasets. At present, the mainstream mapping method is mapping each interface address description datasets to the right information group manually according to the corresponding interface Chinese description texts in the SCD file. Therefore, the classification accuracy of Chinese description texts directly determines the mapping accuracy of these configuration datasets. In high-voltage and large-scale substations, there are many IEDs, and the workload of manual classification is extremely large. For example, a 500kV substation SCD file contains nearly 300 intelligent secondary devices information, and there are tens of thousands of Chinese description texts, which makes the configuration time last more than a month. Besides, the problem of datasets automatic mapping is that there are differences in Chinese description texts. Relative standards have semi-structured constraints on Chinese description texts of different IEDs, but there are still problems in semantic law recognition. In order to meet the high-precision requirements of intelligent recorder automatic mapping task and strengthen its semantic analysis ability of semi-structured text, it is necessary to conduct deep semantic learning for Chinese description text of a large number of IEDs data output interface, and then design the automatic classification mapping unit in the recorder.

Chinese description texts classification consists of four steps: words segmentation and their numerical representation, features extraction and classification. Words segmentation method usually relies on the dictionary constructed with more than 10000 words and uses toolkits to divide them. Traditional numerical representation method usually relies on relevant algorithms to map text words to the numerical space [1], [2]. Traditional features extraction methods rely on eigenvalue function to screen eigenvalues as features [3]–[5]. Traditional classification models include Decision Tree [6], [7], Bayesian classifier [8], [9], Support Vector Machine [10], [11], etc. However, the traditional numerical representation method has two major problems: semantic gap and dimension explosion; the traditional feature extraction method has poor ability to identify typical features; and the traditional classification models relies heavily on specific tasks, and the text association relationship processing is rough [12].

Convolutional Neural Network (CNN) is a typical structure of deep learning frameworks which integrates features extraction and classification as a single step. It can use multiple convolution kernels to collect feature values in different regions from input matrices. Feature vectors representing key semantics are outputted at the top layer of the network and then classified. F. Xu. et al introduce the CNN to the Chinese text sentiment analysis, and the performance of model semantic recognition is better than that of traditional machine learning model [13]. B. Gu. et al introduce the CNN model to human action recognition task, which effectively improves the accuracy of recognition result under different conditions and the efficiency of classification process [14].

B. Zhao and Z. Yang introduce CNN models which are optimized by attention mechanism and crisscross algorithm to the process of power grid load forecasting. Models have small prediction error and great application potential [15], [16]. In a word, Convolutional Neural Network has excellent performance in feature extraction and classification prediction tasks, which can realize Chinese texts deep learning.

This paper presents an automatic mapping method of intelligent recorder configuration datasets based on the DCNN (Dynamic Convolutional Neural Network). Firstly, the SCD file is used to obtain the address description datasets of the IEDs data output interface and the corresponding Chinese description texts. Then, the text representation model Word2vec which is constructed based on BP neural network is introduced to generate the word vectors reflecting the word association relationship. Then, Different DCNN models with different key hyper-parameters are constructed and used to classify the datasets in order to find the optimal model which has the best ability of feature recognition and text classification. Finally, the intelligent recorder configuration datasets are automatically mapped based on the description texts classification result.

II. PREPROCESSING OF INTELLIGENT RECORDER CONFIGURATION DATASETS

Preprocessing of intelligent recorder configuration datasets includes obtaining datasets about IEDs output interface address description information from SCD file, matching the interface address with its Chinese description texts, and representing Chinese texts.

A. OBTAINING CONFIGURATION DATASETS

The configuration datasets of the intelligent recorder are the address description datasets of IEDs data output interface. They are obtained by analyzing the SCD file, which is created by the extensible markup language XML, including nodes, sub nodes, node attributes and their values. Under each IED node, there are sub nodes including LDevice (logical device), LN0 (logical node zero), and FCDA (function constraint data attribute) which contains index information of IEDs interface Chinese description text. A single output interface address could be described by combining a set of FCDA attributes values. At the same time, the FCDA attributes values can also be used as the basis for obtaining the desc attribute value of the sub node DOI (object instance) in the sub node LN (logical node), which is the Chinese description text of the data output interface. Xml.etree.ElementTree [17] module is a commonly used XML file analyzing tool in Python. It built an “Element Tree” to describe different nodes and their affiliation in the original XML file, and the most superior node is called “Root”. All IEDs output interface address description datasets and Chinese description texts can be obtained by traversing all nodes and attributes values, and these two kinds of information could be matched one by one to ensure that the address information mapping result is

the same as the text classification result. Pseudocodes of the obtainment algorithm are shown below.

Algorithm 1 Obtainment of IEDs Output Interfaces Address Description Datasets and Chinese Descriptions (for One Interface)

Parameters:

W_{ij} : 1×7 Variable to storage the j th output interfaces address description datasets of the i th IED and its Chinese description text

Inputs:

The SCD file of an intelligent substation

BEGIN

Initialize W_{ij} with random values

Use the xml.etree.ElementTree module to analyze the SCD file

Root = xml.etree.ElementTree.parse("SCD.xml").getroot()

LDevice_nodes = Root.findall(".//IED/Accesspoint/Server/LDevice")

FCDA_nodes = LDevice_nodes.findall(".//LN0/DataSet/FCDA")

$W_{ij}(1:6) \leftarrow$ FCDA_nodes(j)[“attributes”] (including lnInst, lnClass, doName, daName, prefix, ldInst).

LN_node = LDevice_nodes.findall(".//LN/[@inst = $W_{ij}(1)$ &&@ lnClass = $W_{ij}(2)$ ”)

DOI_node = LN_nodes.findall(".//DOI/[@ doName = $W_{ij}(3)$ ”)

If DOI_node.findall(".//DAI/[@daName= $W_{ij}(4)$ ”) == “True” **do**

$W_{ij}(7) \leftarrow$ DOI_node[“desc”]

Return W_{ij}

B. REPRESENTATION OF CHINESE DESCRIPTION TEXTS

1) FEATURES OF CHINESE DESCRIPTION TEXTS

Compared with common texts, IEDs data output interface Chinese description texts usually involves electrical terminology, such as “远方修改定值软压板 (remote modification of fixed value soft strap)”, “远方操作硬压板 (remote operation hard strap)” and so on. In the words segmentation stage, it is easy to have the false words segmentation result, which leads to the false representation of word vectors. Therefore, the word representation model in this paper will introduce user-defined terms into the classification Python unit “jieba” [18] to improve the accuracy of word vectors representation. At the same time, the text is often the mixture of Chinese and English, such as “切换刀闸位置无效 (the switch position of PT switch is invalid)”, and there are some expression differences in texts which may suggest similar information, such as “链路 3GOOSE 接收 A 网断链 (GOOSE receiving network A of link 3 is disconnected)”, “号板卡光口 F 光强越下限告警 (the GOOSE receiving process of GOCEB1 within the process layer network A is

interrupted)”, when using the expert system to carry out automatic mapping, it is difficult to control the completeness of reasoning rules, resulting in the rules missing or redundant. In this model, the above characteristics are considered in the text representation stage to ensure the accuracy of the Chinese text numerical representation vectors.

2) THE CHINESE DESCRIPTION TEXTS REPRESENTATION MODEL BASED ON Word2vec

Before the description texts are input into the classification model, it needs to be encoded as a word vector matrix. The traditional word representation model BOW (Bag of Words) assigns number ‘1’ to a single dimension in the word vector according to the index value of the sample word in the dictionary and assigns number ‘0’ to other dimensions. Because different word vectors are orthogonal to each other, each word exists in isolation. It is difficult for vectors to represent the semantic relationship between words. In addition, a BOW word vector is high-dimensional and sparse, which makes the classification process time-consuming and inefficient.

Word2vec text representation model [19], [20] (including Continuous BOW and Skip-Gram two kinds of forms), which is built based on BP neural network, can effectively reduce the dimension and sparsity of word vectors and improve the vectors ability to represent semantic relationship of the original words. The model sets local text analyzing window and assumes that the context word vector in the window is known Then it regards maximizing the occurrence probability of the central word as the Word2vec model training target, in order to realize the mapping of words in the semantic space and obtain the dense static word vectors after dimension reduction.

In this paper, Continuous BOW model is used to represent the words in IEDs data output interface Chinese description texts, and negative sampling method is used to optimize the weight-updating process. With the sliding of the text analysis window, each word can appear as the central word and the context word. Therefore, Continuous BOW model contains the context vector matrix and the central word vector matrix. These two matrices are randomly initialized. When a word appears as the context word, the model inputs its BOW word vector to obtain the context vector. Then the word appear as the central word and the input value of the output layer is obtained by multiplying the central word vector of the word and the context word vector of the nearby word. Through training the Continuous BOW model, the parameters of the context word vector matrix and the central word vector matrix can be updated. Finally, the word vector of each word is the average value of its context word vector and central word vector. The final numerical representation result of a single text is a matrix which is constructed by connecting word vectors in sequence. The schematic diagram of Continuous BOW is shown in Figure 1, in which $w_1 \dots w_c$ are different words included in a sample text and w_i is the central word, W_1 is the context vector matrix and W_2 is the central word

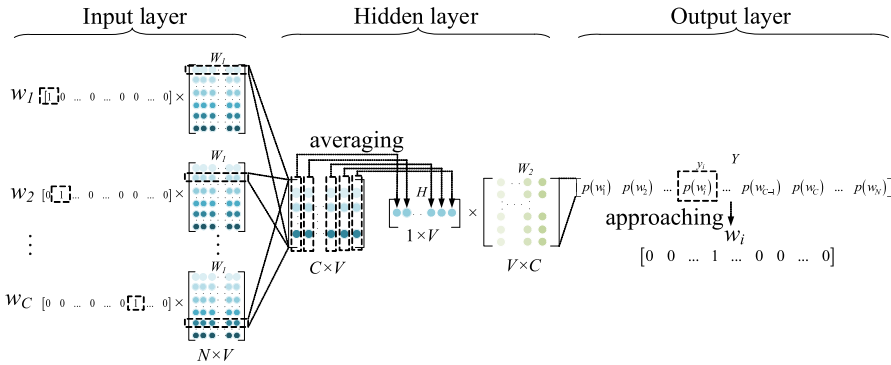


FIGURE 1. Schematic diagram of Continuous BOW model.

vector matrix, $p(w_1) \dots p(w_c)$ are words probabilities the model calculates, H is the output of the hidden layer. Y is the output of the Continuous BOW model, N is the number of words included in the dictionary, C is the number of words included in a sample text and V is the dimension of word numerical representation vector. The method to calculate y_i is shown in (1), in which v_{1w_j} is the context word vector of w_j and $v_{2w_j}v_{2w_k}$ are central word vectors of w_i and w_k

$$y_i = p(w_i) = \frac{\sum_{j=1}^C \exp(v_{2w_i}^T v_{1w_j})}{\sum_{j=1}^C \sum_{k=1}^N \exp(v_{2w_k}^T v_{1w_j})} \quad (1)$$

The calculation process from the hidden layer to the output layer includes solving the occurrence probability of all words and finding the maximum value which leads to the calculation amount complex and time-consuming. In order to improve the efficiency, the model introduces the negative sampling method [20], which sets one context word included in the analyzing window as the positive sample, and negative samples choose five other word vectors not included in the analyzing window according to the probability, so that the number of matrices element values to be updated at a single time is reduced to 5% of the original number, Thus, the calculation time is greatly reduced. The selection probability is shown in (2): where, w_j is the positive sample and $f(w_j)$ is the frequency of negative sample words w_j 's occurrence. The loss function is shown in (3), in which W_{neg} is a set of words not to be solved and σ is the Sigmoid function.

$$p(w_j) = \frac{f(w_j)^{\frac{3}{4}}}{\sum_{k=1}^N (f(w_k)^{\frac{3}{4}})} \quad (2)$$

$$E = -\log \sigma(v_{2w_i}^T h) - \sum_{w_j \in W_{neg}} \log \sigma(-v_{2w_j}^T h) \quad (3)$$

$$h = \frac{1}{C} \left(\sum_{k=1}^V v_{1w_k} \right)^T \quad (4)$$

III. AUTOMATIC MAPPING OF INTELLIGENT RECORDER CONFIGURATION DATASETS

A. AUTOMATIC MAPPING OF CONFIGURATION DATASETS

1) THE CONFIGURATION DATASETS MAPPING MODEL BASED ON THE DYNAMIC CONVOLUTION NEURAL NETWORK

Convolution neural network (CNN) is the most powerful deep learning network framework for local feature mining in computer image processing tasks. In the field of text classification, when the text is represented as a word vector, the vector value can be seen as the image gray value, and the word vector matrix can be processed as a one-dimensional gray image, which is input to CNN for local semantic feature analysis, text key numerical value selection and classification. Traditional text classification CNN includes convolutional layer, pooling layer and full connection layer, in which the pooling layer usually adopts the maximum pooling method to reduce the dimension of input vectors, and the output vector only uses a single key feature value to describe the input vector, ignoring the correlation between the omitted value of the vector and the feature value. Thus, the semantic learning ability of the traditional CNN is weak.

In view of this, this paper introduces the DCNN (Dynamic Convolution Neural Network) as the Chinese text classifier, in which dynamic K -max pooling layer is adopted to reduce the dimension of the input vectors, and key feature values of different numbers are extracted according to the length of each text sentence, in order to retain the relative position between the words and the word order information in the original text. In addition, the wide convolutional layer in the DCNN model can not only effectively capture the local feature value but also the edge information of the text more comprehensively. Generally speaking, DCNN structure includes four kinds of network layers: wide convolutional layer, K -max pooling layer and dynamic K -max pooling layer, folding layer and full connection layer. The structure schematic diagram is shown in Figure 2. The input text is “the GOOSE receiving process of GOCB1 within the process layer network A is interrupted”, which has seven words in Chinese.

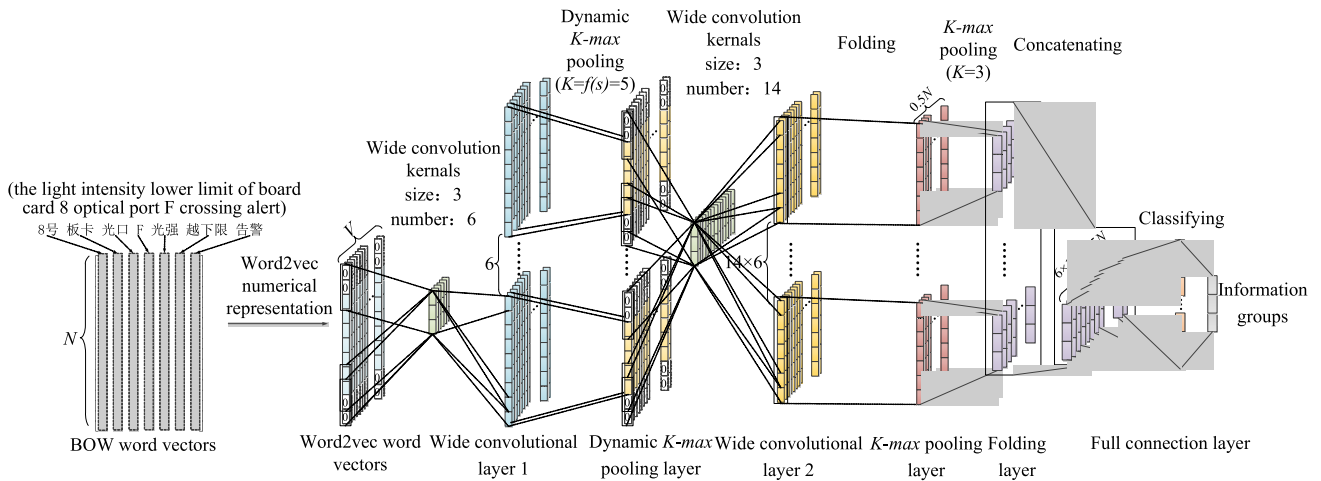


FIGURE 2. The structure of the Dynamic Convolution Neural Network.

In the wide convolutional layer, $1 \times m$ -dimension wide convolution kernel is used to convolute vectors from different dimensions of the word vectors matrix. m is the convolution dimension. w is a single dimension value of the convolution kernel, and x is a single dimension value of a word vector. o is the output of a single dimension value after convolution, and b is the model offset. the calculation process is shown in (5). Different from the common narrow convolution kernels, when processing vectors with wide convolution kernels, the edge of vectors is usually filled with zeros to retain the edge semantic information.

$$o = f(w \cdot x + b) = \max(0, w \cdot x + b) \quad (5)$$

The K -max pooling layer and the dynamic K -max pooling layer are used to filter the key feature values. The K -max pooling layer selects top K largest feature values from each dimension of the input vector, and the selected values retain the typical semantic and order information of the input vector to the greatest extent. When the order of feature values increases, the processing of fixed number maximum pooling will easily lead to the redundancy of semantic information selection. To ensure that the number of values of pooling layer can be adjusted adaptively to the length of input text sentence, the dynamic K -max pooling layer sets K as a function of text sentence length and network depth, as is shown in (6). Where n is the input sentence length, l is the sorting number of the current layer in all convolutional layers, L is the number of convolutional layers, and K_{top} is the K value of the top-level pooling operation.

$$K = \max\left(K_{top}, \left\lceil \frac{L-l}{L}n \right\rceil\right) \quad (6)$$

In the folding layer, the adjacent two-dimensional vector values of the input matrix are summed and spliced. In the previous calculation process, the convolution operation is only carried out for each dimension of the input vector and different dimensions are independent of each other. The folding layer can map the relationship between adjacent

dimensions, and the input vector dimension can be reduced by half. Finally, the full connection layer realizes the text classification.

The DCNN classification prediction results will be imported into the Python dictionary that matches the IEDs output interface Chinese description texts and their address description datasets, which makes preparation for the datasets mapping into the configuration file of the intelligent recorder.

2) INPUTTING CONFIGURATION DATASETS INTO THE INTELLIGENT RECORDER CONFIGURATION FILE

The configuration file of the intelligent recorder is also written in XML language. Each IED_node contains a_node set constructed by three major information groups and their sub information groups. These_nodes also contain a blank address description sub_node to be filled by the address description datasets of the IED output interface. This paper uses the xml.etree.ElementTree module in Python to read the classification results of Chinese description texts before inputting address information, and loop each information group of IED_node to find the target category according to the classification results. Then the model automatically input the interface address description datasets to the address description sub_node, completing the automatic mapping of the datasets. When the recorder is put into use, it can automatically find the address of each IED data output interface by analyzing address description sub_node contained in each information group of its configuration file, so as to accurately monitor IED operation information. Pseudocodes of algorithm inputting datasets into the configuration file are shown below.

IV. EXPERIMENTS

A. INFORMATION ABOUT SAMPLES

This paper selects 2500 classified Chinese description texts of IEDs data output interface recorded in 110kV and 500kV substations in Hebei Province as analysis samples. The description text and information groups which the

Algorithm 2 Inputting Interface Address Description Datasets Into the Intelligent Recorder Configuration File

Parameters:

W_{ij} : 1×8 Variable, in which $W_{ij}(1:7)$ are values of attributes including daName, doName, lnInst, lnClass, prefix, ldInst from the FCDA_node in SCD file, $W_{ij}(8)$ is the Chinese description text (e.g. “8重合闸(the light intensity lower limit of board card 8 optical port F crossing alert)”), $W_{ij}(7)$ is the classification result (Online_Monnitoring_Information)

FCDA: a_node in the configuration file of an intelligent recorder, FCDA (1:6) are values of attributes including lnInst, lnClass, doName, daName, prefix, ldInst; FCDA (7) is the Chinese description text. At the beginning all values are vacant.

Inputs:

The configuration file of an intelligent recorder

BEGIN

Read $W_{ij}(8)$ result(=Online_Monnitoring_Information)

Use the xml.etree.ElementTree module to analyze the Intelligent_Recorder_Configuration_File

Root = xml.etree.ElementTree.parse

(“Intelligent_Recorder_Configuration_File.xml”).
getroot()

FCDA_node = Root.findall

(“./IED/Information_Groups/Status_Monitoring_Information/Online_Monnitoring_Information/FCDA”)

FCDA_node[“attributes”] ← $W_{ij}(1:7)$

TABLE 1. A part of samples descriptions.

information group	Chinese description texts
the functional strap information group	高压侧相间后备保护软压板 (the inter phase backup protection soft strap from high voltage side), 远方修改定值软压板 (remote modification setting soft strap)
the optical brand alarming information group	装置内部温度 (the internal temperature of the device), 正母分段汇控柜湿度 (the humidity of the master bus sectional control cabinet)
the online monitoring information group	8号板卡光口F光强越下限告警 (the light intensity lower limit of board card 8 optical port F crossing alert), 时间跳变检测状态 (the detection status of time jump)

datasets belong to are complete. In this paper, samples are divided into the train set, validation set and the test set in a ratio of 6:2:2. The model uses the train set and validation set to update model parameters and the test set to evaluate model performance in each epoch. In this paper, three typical subcategories are selected as the target class to which the sample datasets belong: the functional strap information group in the strap information group, the optical brand alarming information group in the alarming information group, and the online monitoring information group in the status monitoring information group. Some text samples are shown in table 1. This model is programmed in Python 3.6.8,

and word vector is constructed by gensim module. DCNN classifier is built by tensorflow toolkit. The CPU of the computer is Intel Core i7-8565u, the main frequency is 1.8GHz, the running memory is 8g, and the capacity of solid-state disk is 256gb.

The metric F_1 is the mostly used evaluation metrics in the common binary classification. Table 2 shows the confusion matrix, which is a basis for defining the metrics F_1 .

TABLE 2. Confusion matrix.

Actual class	Predicted class	
	Positive	Negative
Positive	True Positive (TP)	False Negative (FN)
Negative	False Positive (FP)	True Negative (TN)

Therefore, the P , R , and F_1 can be defined as:

$$P = \frac{TP}{TP + FP} \quad R = \frac{TP}{TP + FN} \quad F_1 = \frac{2 \times P \times R}{P + R} \quad (7)$$

However, the texts can be divided into three groups (i.e. the functional strap information group / the optical brand alarming information group / the online monitoring information group), so it is a three-class problem. Hence, $macro-F_1(MF_1)$ is adopted to measure the average performance of whole groups. MF_1 is calculated by averaging F_1 of all groups.

$$MF_1 = \frac{2 \times MP \times MR}{MP + MR} \quad (8)$$

where, MP and MR are macro- P and macro- R respectively. They are defined in (9).

$$\begin{cases} MP = \sum_{m=0}^2 P_m / M \\ MR = \sum_{m=0}^2 R_m / M \end{cases} \quad (9)$$

where, P_m and R_m are the calculation results corresponding to the class l_m ($m = 0, 1, 2$).

B. CALCULATION AND ANALYSIS OF THE EXAMPLE

1) EVALUATION OF Word2vec REPRESENTATION ABILITY

The Word2vec model could represent words semantic relationship in the original text. In order to prove that, this paper introduces the Principal component analysis (PCA) method, which is a common algorithm capable of reducing the dimension of high-dimensional vectors and converting them to low dimensional coordinates in another vector space. After the Word2vec vectors of different words being processed by the PCA method, the words semantic relationship could be measured by the distances between different coordinates in the three-dimensional semantic space, and the distances are used to evaluate the Word2vec representation ability. If the distances between words with close semantic relationship are much shorter than that of words with sparse semantic relationship, the Word2vec representation ability is proved to be strong.

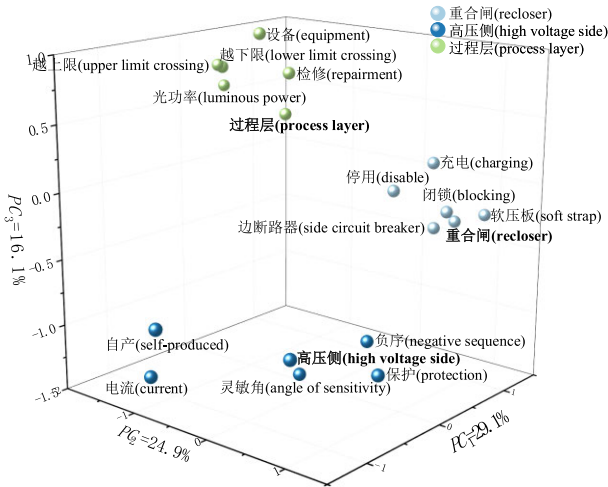


FIGURE 3. Sample words in the three-dimensional semantic space.

This paper chooses 3 different words including “高压侧 (recloser)”, “过程层 (high voltage side)” and “过程层 (process layer)” coming from different description texts as central words. To make the relationship between different words more obvious, five most closely related words of each central word are also selected as background words. In total, 18 Word2vec word vectors are chosen to be processed by the PCA method and mapped into the three-dimensional semantic space. The results are shown in Figure 3. Since each central word only appears with its background words together in a single description text instead of the other 12 words, the distances between each central word and their background words are significantly shorter than that of other words, which reveals the fact that each central word has closer semantic relationship with its background words. Finally, three clusters with three central words as their centers could be obviously observed in the three-dimensional semantic space. Equation (10) defines the method to calculate the relativity between words coordinates $w_l(x_l, y_l, z_l)$ and $w_m(x_m, y_m, z_m)$, in which L_{max} is the largest distance between words in the three-dimensional semantic space. The bar charts of relativities calculation results are shown in the Figure 4,

in which three groups of bars representing relativities between three groups of background words closely related to the central words and central words themselves.

$$Relativity_{lm} = \frac{\sqrt{(x_l - x_m)^2 + (y_l - y_m)^2 + (z_l - z_m)^2}}{L_{max}} \quad (10)$$

2) EVALUATION AND OPTIMIZATION OF DCNN CLASSIFICATION ABILITY

The hyper-parameters are set to fixed values before model training. Usually, hyper-parameters need to be optimal to improve the performance and effectiveness of learning. There are four key hyper-parameters closely related to the ability of the model. The learning rate and batch size directly determine the speed and range of model parameters updating within a single batch. The convolution kernel size decides the size of key semantic feature value, the more features considered in a single time, the larger the convolution kernel size is. The dropout ratio of full connection layer directly affects the generalization ability of the model. This paper adjusts hyper-parameters referring to the sentiment prediction sentiment prediction project in movie reviews [21]. Firstly, the convolution kernel size is set to 3 and the dropout ratio dp is 0.5, and the contrast experiments between models with different learning rates and batch sizes shows that the model whose learning rate lr equals 0.01 and batch size bs equals 64 has the highest test set MF_1 . The experiments results are shown in Figure 5.

Secondly, contrast experiments between models with different kernel sizes and batch sizes are conducted to choose the kernel size and dropout ratio. Models with three groups of different convolution kernel sizes are set according to the sample text length and the commonly used kernel sizes of text convolution network. Three groups of dropout ratios are also set up, in which 0.8 indicates that the reserved value is less, which may lead to the loss of key semantics, and 0.2 is close to complete retention, which results in poor generalization performance and excessive dependence on local features. The experiment results are shown in Figure 6. According to the result, when the kernel size equals 4 and the dropout ration equals 0.5. the MF_1 of test set reaches the highest

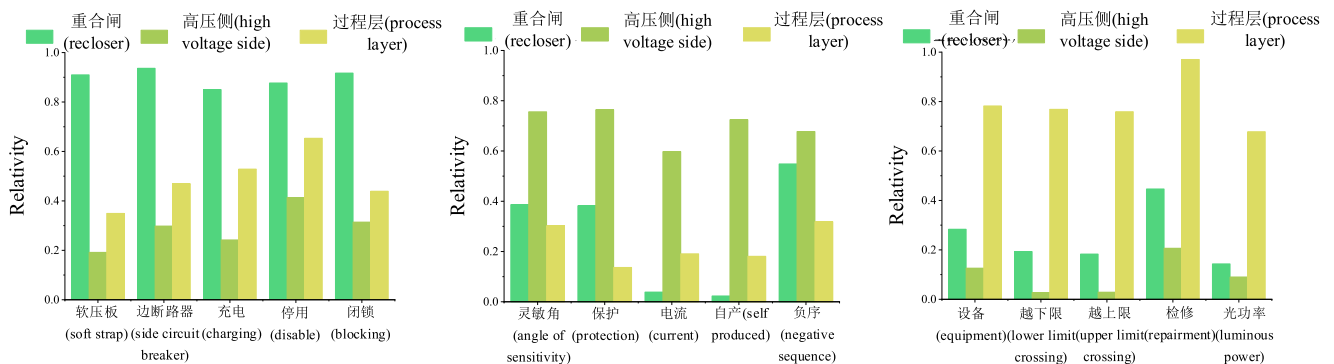


FIGURE 4. Relativities between three central words and their five most closely related words.

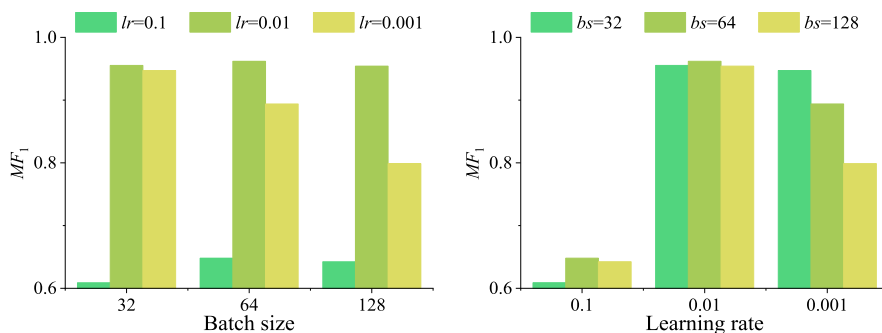


FIGURE 5. Relationships between MF_1 and hyper-parameters (learning rate and batch size).

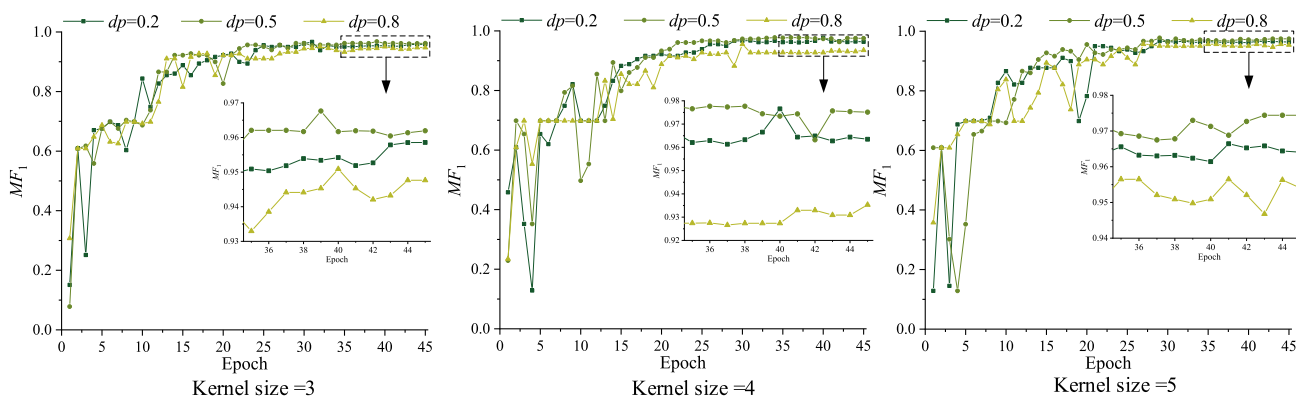


FIGURE 6. Relationships between MF_1 and hyper-parameters (convolution kernel size and dropout rate).

TABLE 3. Hyper-parameters settings.

	Value
convolution kernel size	4
number of convolution kernels in the first convolutional layer	6
number of convolution kernels in the second convolutional layer	14
K_{top} value of the K-max layer	4
dimension of word vector	64
learning rate	0.01
number of neurons in the full connection layer	128
dropout ratio	0.5
maximum training epochs	45
batch size	64

value 97.52%. After that, enlarging the kernel size does not effectively increase the MF_1 . The final hyper parameters of the model are shown in Table 3.

In order to reflect the superior performance of the automatic mapping model DCNN, this paper selects four typical traditional classification models and two kinds of shallow neural network classifiers for comparison. The training set is used to update model parameters, and the evaluation standard is test set MF_1 . Four kinds of traditional

TABLE 4. Hyper-parameters settings.

	test set MF_1 .
SVM+BOW	78.32%
kNN+BOW	89.48%
CART+BOW	90.35%
NaiveBayes+BOW	67.14%
BPNN+Word2vec	93.41%
CNN+Word2vec	93.53%
DCNN+Word2vec	97.52%

models include Support Vector Machine (SVM), K nearest neighbor algorithm (KNN), cart Decision Tree and Naive Bayesian algorithm, and the text numerical representation model is BOW; The shallow neural network includes BPNN and CNN, and the text numerical representation model is Word2vec. The experimental results are shown in table 4. It can be seen that the MF_1 of the classification model based on Word2vec and DCNN is significantly higher than other models, which proves that the classification ability of the DCNN is excellent.

V. CONCLUSION AND FUTURE WORK

This paper considers that when mapping the IEDs data output interface address description datasets to the different information groups according to Chinese description texts, the manual mapping workload is large and automatic mapping method could have high error rate due to the

complexity of the data and the semi-structure feature of description texts. The paper proposes a self-mapping method of address description datasets based on the Dynamic Convolution Neural Network. After obtaining the Chinese description texts and representing them through BOW model, the Word2vec is used to reduce the dimension of the BOW word vectors, map the semantic relationship between different words and therefore calculate the final word vectors. Then, the DCNN model is used to select local semantic feature values of word vectors, identify and extract key semantics for classification, which greatly improves the classification accuracy. The classification time is far less than the manual configuration work time, and the comprehensive evaluation index $macro-F_1$ of test set samples can reach more than 97%, which is significantly higher than that of the traditional classification model and shallow neural network. Through providing the mapping result for the configuration datasets input into the intelligent recorder configuration file, the DCNN model effectively improves the accuracy of automatic mapping of recorder configuration datasets.

REFERENCES

- [1] Y. P. Liu, Z. Xu, and G. Li, "Review on applications of artificial intelligence driven data analysis technology in condition based maintenance of power transformers," *High Voltage Eng.*, vol. 45, no. 2, pp. 337–348, Feb. 2019.
- [2] X. M. Du, J. F. Qin, S. Y. Guo, and D. F. Yan, "Text mining of typical defects in power equipment," *High Voltage Eng.*, vol. 44, no. 4, pp. 1078–1084, Apr. 2018.
- [3] X. X. Li, "Research of text categorization based on improved maximum entropy algorithm," *Comput. Sci.*, vol. 39, no. 6, pp. 210–212, Jun. 2012.
- [4] J. Wang, L. Jin, and K. Sun, "Chinese text categorization based on evolutionary hypernetwork," *J. Jiangsu Univ. (Natural Sci. Ed.)*, vol. 34, no. 2, pp. 196–201, Apr. 2013.
- [5] Y. X. Zhu, "Text classification algorithm research based on clustering and neural network," *Appl. Res. Comput.*, vol. 29, no. 1, pp. 155–157, Jan. 2012.
- [6] S. R. Safavian and D. Landgrebe, "A survey of decision tree classifier methodology," *IEEE Trans. Syst., Man, Cybern.*, vol. 21, no. 3, pp. 660–674, Jun. 1991.
- [7] Z. C. Jia, Q. Han, Y. Li, Y. Yang, and X. Xing, "Prediction of Web services reliability based on decision tree classification method," *Comput., Mater. Continua*, vol. 63, no. 3, pp. 1221–1235, Apr. 2020.
- [8] G. Santafe, J. A. Lozano, and P. Larranaga, "Bayesian model averaging of naive Bayes for clustering," *IEEE Trans. Syst., Man, Cybern. B. Cybern.*, vol. 36, no. 5, pp. 1149–1161, Oct. 2006.
- [9] T. Shen, "Improve computer visualization of architecture based on the Bayesian network," *Comput., Mater. Continua*, vol. 58, no. 2, pp. 307–318, Feb. 2019.
- [10] C.-W. Hsu and C.-J. Lin, "A comparison of methods for multiclass support vector machines," *IEEE Trans. Neural Netw.*, vol. 13, no. 2, pp. 415–425, Mar. 2002.
- [11] F. Bi, X. Fu, W. Chen, W. Fang, X. Miao, and B. Assefa, "Fire detection method based on improved fruit fly optimization-based SVM," *Comput., Mater. Continua*, vol. 62, no. 1, pp. 199–216, 2020.
- [12] Y. W. Jiang, L. Li, Z. Li, C. Su, G. Wang, and X. Xiao, "An information mining method of power transformer operation and maintenance texts based on deep semantic learning," *Proc. CSEE*, vol. 39, no. 14, pp. 4162–4172, 2019.
- [13] F. Xu, X. Zhang, Z. Xin, and A. Yang, "Investigation on the Chinese text sentiment analysis based on convolutional neural networks in deep learning," *Comput., Mater. Continua*, vol. 63, no. 1, pp. 243–262, Apr. 2020.
- [14] B. Gu, W. Xiong, and Z. Bai, "Human action recognition based on supervised class-specific dictionary learning with deep convolutional neural network features," *Comput., Mater. Continua*, vol. 58, no. 3, pp. 697–709, Mar. 2019.
- [15] B. Zhao, Z. P. Wang, W. J. Ji, X. Gao, and X. B. Li, "CNN-GRU short-term power load forecasting method based on attention mechanism," *Power Syst. Technol.*, vol. 43, no. 12, pp. 4370–4376, Dec. 2019.
- [16] L. Yang, L. Q. Zhong, and Y. Hao, "Research on two-dimensional combined short-term load forecasting method based on convolutional neural network and vertical and horizontal cross algorithm," *Mod. Inf. Technol.*, vol. 3, no. 4, pp. 160–162, Feb. 2019.
- [17] The Python Software Foundation. (2018). *Xml.etree.ElementTree-The ElementTree XML API, Python 3.6.8, Documentation*. [Online]. Available: <https://docs.python.org/release/3.6.8/library/xml.etree.elementtree.html#module-xml.etree.ElementTree>
- [18] J. Y. Sun. (Jan. 20, 2020). *Jieba (Chinese for 'to Stutter') Chinese Text Segmentation: Built to be the Best Python Chinese Word Segmentation Module, Python Package Version 0.42.1*. [Online]. Available: <https://github.com/fxsjy/jieba>
- [19] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," in *Proc. 1st Int. Conf. Learn. Represent.*, Scottsdale, AZ, USA, 2013, pp. 1–13.
- [20] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," in *Proc. Adv. Neural Inf. Process. Syst.*, Lake Tahoe, NV, USA, 2013, pp. 3111–3119.
- [21] N. Kalchbrenner, E. Grefenstette, and P. Blunsom, "A convolutional neural network for modelling sentences," presented at the 52nd Annu. Meeting Assoc. Comput. Linguistics, 2014. [Online]. Available: <https://arxiv.org/abs/1404.2188>



JIANGBO REN was born in Shijiazhuang, Hebei, China, in 1979. He received the B.S. and M.S. degrees in electrical engineering from the Harbin Institute of Technology, in 2003.

From 2004 to 2020, he was an Engineer with State Grid Hebei Electrical Power Company Research Institute, Shijiazhuang, Hebei. His research interest includes relay protection control technology.



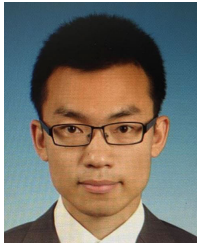
TIECHENG LI was born in Shijiazhuang, Hebei, China, in 1980. He received the M.S. degree in electrical engineering from North China Electric Power University, Baoding, in 2007.

From 2007 to 2020, he was an Engineer with State Grid Hebei Electrical Power Company Research Institute, Shijiazhuang, Hebei. His research interest includes relay protection control technology.



SHAOBO GEN was born in Tangshan, Hebei, China, in 1986. He received the M.S. degree in electrical engineering from North China Electric Power University, Baoding, in 2009.

From 2009 to 2020, he was an Engineer with State Grid Hebei Electrical Power Company, Shijiazhuang, Hebei. His research interests include electrical engineering and new energy operation control.



QINQUAN LIU was born in Baoding, Hebei, China, in 1988. He received the M.S. degree in electrical engineering from North China Electric Power University, Beijing, in 2010.

From 2010 to 2020, he was an Engineer with State Grid Hebei Electrical Power Company Research Institute, Shijiazhuang, Hebei. His research interests include operation and management of power grid dispatching control.



ZHIHUA WANG was born in Wuhan, Hubei, China, in 1970. He received the M.S. degree in electrical engineering from North China Electric Power University, Beijing, in 2012.

From 1990 to 2020, he was an Engineer with Wuhan Kemov Electrical Company Ltd., Wuhan. His research interest includes relay protection operation technology.



YAKUN HE was born in Baoding, Hebei, China, in 1986. He received the M.S. degree in electrical engineering from North China Electric Power University, Beijing, in 2017.

From 2017 to 2020, he was an Engineer with State Grid Hebei Electrical Power Company, Shijiazhuang, Hebei. His research interest includes operation and management of power grid dispatching control.



DAMING ZHOU was born in Changsha, Hunan, China, in 1997. He received the M.S. degree in electrical engineering from the Changsha University of Science and Technology, in 2019.

From 2019 to 2020, he was an Engineer with Wuhan Kemov Electrical Company Ltd., Wuhan, Hubei, China. His research interest includes smart grid data mining.

...