Optimal Design of Retrieval System Based on The Deep Learning

Qiyu Han [a], Panqing Wang, Jianlong Li

Ordnance Engineering College, 050003, ShiJia zhuang, China;

[a]505151776@qq.com

Abstract

This paper briefly describes the current development of the deep neural network (DNN). On basis of traditional B-P neural network and experiments, the most suitable number of hidden layers for equipment textual information mining is chosen to establish a deep neural network. The S-type function in the original B-P neural network is replaced by ReLU integral linear function so as to refine the convergence rate, reduce the node number in the whole network and further improve retrieval system.

Keywords

Data mining, deep neural network, text mining, information retrieval, fault prediction.

1. Introduction

The US forces have been in information wars of various scales for many times at the beginning of the 21st century around many countries and regions, and the United States Department of Defense also successively issued a series of key projects and investment areas in national defense scientific researches, which focus on strengthening promotions of information and technology development including <Network-Centric Warfare>[6] and <Defense Technology Strategy>. The realization of the potential knowledge and the laws of equipment information in the massive equipment data makes it a fact that equipment information resources can be fully utilized, which has become a key step in our military information construction. So the military has been absorbing itself in enhancing the promotions of information technology and the training of professionals in this field. At the same time, the data mining technology such as the structured data and semi-structured data has been relatively mature in the equipment information period[7], but when it comes to the data stored in various types of equipment inspection report with the form of natural language, it cannot be fully utilized. Therefore, the text data of the military equipment information will be considered as a basis in this paper. The method of text mining based on deep neural network is preferred and it is popular in the world to discover the knowledge and potential data rules hidden in the equipment detection report. Besides, it can be used to optimize the equipment information retrieval and predict the equipment failure. The purpose of this paper is to enhance the full use the data of the military equipment information and satisfy the demands of rapid development in information technology military equipment. Meanwhile, it is enable to make the military more perfect in predicting equipment failure, which helps to make better and faster decisions in the field of equipment.

2. Research on Deep Neural Network Technology

2.1 The Development of Deep Neural Network

The artificial neural network refers to a network system established by artificial methods of simulating functions and structures in human beings' brain system through numerous utilization of processing components[4]. The artificial neural networks can be also understood as a kind of structural mode in which the distributed parallel information is processed by the processing units (artificial neurons, etc). The earliest proposal of this artificial neural network was given by

W.Mculloch and W.Pitts in the year of 1943, and they proposed the M-P model for the first time. Rosenblatt proposed a perceptual machine model in 1958, which is a relatively simple neural network model. Since the study from scholars including J.J. Hopfield explored plenty of potentials in neural networks, after a short period of stagnation, the study of artificial neural networks also met with its short boom in the 1980s. But the artificial neural network still has some limits, hence it inevitably went into declines once again. With the development of the times, however, the artificial intelligence begun to occupy the dominant position. Under the rapid development of artificial intelligence, the machine learning method of deep neural network based on human brains' thinking, which is proposed by researchers like Hinton in 2006, makes the neural network become the whole world's hot spot once again. This paper firstly proposed the definition of deep belief network[2], the basic purpose is to supervise and train a number of restricted Boltzmann machines, besides, it takes the completed weight as the initial weight and adds classifiers for recognition finally. The problems caused by those increasing layers in the artificial neural network model, including phenomenons that the convergence speed is getting slower and the network model is easier to trapped in local optimum, have been solved exactly with the help of network.

With the further use of deep learning in all fields, the deep learning has made great breakthroughs and progress. As for the textual processing area, the deep learning also owns some advantages that the machine learning can not match in the past, which means it can give more conversions to simple documents through the deep network and form a multi-layer structure. This allows us to attain a qualitative leap in dealing with the natural language, at the same time, the utilization of deep learning in the document information explores the information of equipment test reports which are restored by texts in a better way. Hence, it provides a more reliable guarantee for handling the textual information in the field of equipment. In the original equipment area, there are methods to deal with the structured data and the semi-structured data. The equipment data can be also fully used and enhanced further by combing former methods and those related to textual information in deep learning, so a more complete systematic structure in predicting the equipment quality can be obtained to provide some better decision-making services for this field.

2.2 Principles of Deep Neural Networks

Similar to the data mining technique, the deep learning also has a variety of definition claims. One of the arguments is that a simple data can be achieved by increasing the number of the middle layers and the corresponding node data in each layer under the basis of the construction of traditional neural networks, which enables to express it as the high latitude of the multiple features to achieve the correct rate of improvement ultimately.

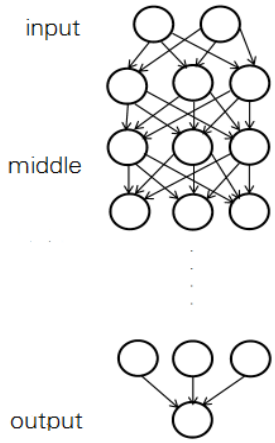

Fig.1 Depth Learning Network Model

As it can be shown in the picture of 1, the model is based on the neural network. We all know that a neural network model with a hidden layer is called a shallow neural network while a neural network model with two or more hidden layers is called the deep neural network, which is also known as the deep learning network model. Compared with the traditional neural network model, the advantage of

the deep learning network model is that the nonlinear transformation of the traditional neural network can be transformed into various forms combined by nonlinear operations under the increase of the deep network, which is the accumulation of hidden layers in the model. In this way, the deep learning model has higher accuracy and better expression ability. The central idea of the deep learning can be summarized as the following points[1] :

1. Deep learning is a multi-level neural network model, which means each layer of neural networks is under the unsupervised learning mechanism in the condition of pre-training.

2. The training of the deep learning is carried out layer by layer, in other words, the output of the prior level is the input of next one.

3. In the process of the deep neural network training, a classifier is needed to cooperate with it, that is to say, the deep learning process is accompanied by supervised learning.

2.3 The textual preprocessing

The basis of the textual data for this paper is based on test report of the military equipment field, which includes not only the cause of failure but also the equipment operator's personal information such as his age, especially his military age and his military appointment. Besides, the operation equipment information such as the time and equipment usage, equipment types, and equipment parameter standards are also included. Moreover, the equipment maintenance information, taking equipment maintenance time as an example, is also needed in this test report. Finally, an expected output result will be achieved through a merging process. For example, if an equipment breaks down, a certain type of maintenance frequency will be checked and an overall record will be summarized including the equipment maintenance consuming time. There is no denying the fact that the output results will be compared with the expected output results, and the output results which meet the conditions will be retained while the invalid data removed.

The preprocessing for the textual information can be broadly divided into the following steps:

(1) Using the word segmentation tool declared by Chinese Academy of Sciences to deal with the equipment failure detection report can delete the stop words from the test report together with the utilization of latest stop-word package which has been processed. It can reduce the interdiction of stop words.

(2) Key words extraction is beneficial to the construction of the deep neural network model. The weight of the keywords plays a crucial role in selecting the hidden layer of the deep learning as well as the input function of the input layer.

(3) The knowledge base should be built. The results of the Chinese segmentation words can not be directly used, and those existing segmentation words should be re-processed in accordance with the features of our military equipment as well as experts' relevant experience in the field of equipment. For example, it is demanded that the semantic heterogeneous words should be covered in the equipment field, that is to say, in the condition that both equipment types and fault types are similar but there are different issues in the expressing ways, a united regulation should be made clearly ahead of time in expressing ways of semantic heterogeneous in segmentation words. In this way, it will reduce the influences of output process due to the semantic heterogeneity, and it will also reduce the intervention of semantic heterogeneous. At the same time, the knowledge base is also classified according to the word segmentation results, which makes it easier to search.

(4) Using the word2vec open-source software developed recently, the text information will be mapped to an R-dimensional vector space, then the complex textual information will be expressed in the form of vector set. The breakthrough of the word2vec software is that the vector operation between keywords can be correspond to semantics. In the process of vector projection, choosing a reasonable dimension as a difference value and making a reasonable projection are both needed.

2.4 The Optimization Design of Equipment Information Retrieval Based on Deep Neural Network

2.4.1 The construction of the multi-layer neural network model

Currently, the B-P neural network is the most widely used as a model of feed-forward neural network [5], meanwhile it is suitable for neural network backward learning from one layer to multi-layer. Therefore, this experiment will be carried out on the basis of the B-P neural network model, in which the initial hidden layer number will be set to 1 and the end to n. As is shown in Figure 2.1, in other words, it can be considered that one hidden layer will be achieved in initial states while n hidden layers in final states according to this model. The number of hidden layers will be selected by experimental comparisons so as to construct the basic model of the deed neural network. Besides, both feed-forward training and reverse propagation training are needed for one training process in B-P neural network. We set the input text keywords to x1, x2, …, xn, and the connection weights to w1, w2, …, wn, after removing stop words, the weight of each keyword is $\frac{1}{w_i}$ (i=1, 2, …, n). For further communication training, the probability of each textual keyword is $\frac{1}{w_1}+\frac{1}{w_2}+\ldots+\frac{1}{w_n}$ through the selection of keywords, and then set the number of keywords for each text as H so that the probability of keywords occurrence can be expressed as $\frac{\frac{1}{w_1}+\frac{1}{w_2}+\cdots+\frac{1}{w_n}}{H}$. For reverse communication training, it is necessary to use the probability $\frac{\frac{1}{w_1}+\frac{1}{w_2}+\cdots+\frac{1}{w_n}}{H}$ of the recurrent keywords as the input through the reverse transmission. In addition, the probability of the keywords which appear in the text but not in the reverse transmission is $\frac{1}{w_i}+\frac{\frac{1}{w_1}+\frac{1}{w_2}+\cdots+\frac{1}{w_n}}{H}$. Finally, a training process of the B-P neural network comes to an end after the feed-forward training and the back propagation training.

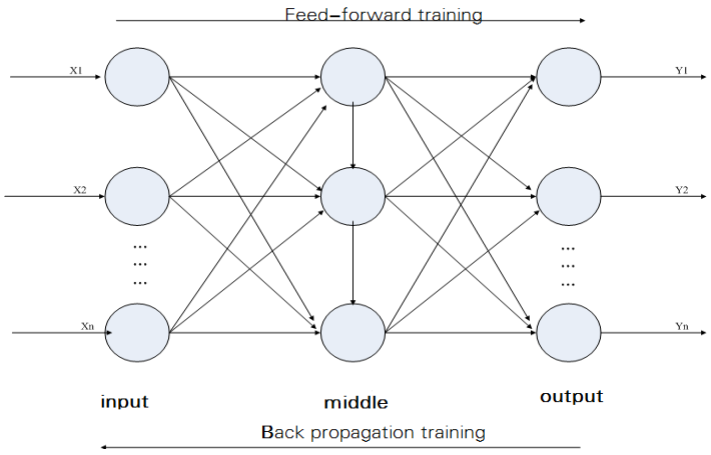

Fig.2 B-P neural network model with one hidden layer

Based on the 1-layer of the neural network model and by increasing the hidden number of middle parts and creating deep neural network model, which is also known as the multi-layer neural network model. As is shown in Figure.2, the deep neural network model which is suitable for the textual information data mining will be achieved in this paper through this experiment.

We proposed a multi-layer neural network model with a hidden layer of neural network as the experimental model. Besides, a more characteristic network model will be selected for convenient explanations in this paper. As is demonstrated in Figure .3 of this model, the layer number is set to 3 as, and five neural networks can be decomposed into two neural network models with one hidden layer. Each model can be regarded as a three-part one including input layer, intermediate node and output layer. In the process of the first network model, the users' search will be considered as an input to extract the keywords through the middle hidden layer connection. Then the keywords will be the input layer in the second network model, connecting the final output textual information through the

middle hidden layer. In the first network model, the middle hidden layer represents a relationship between the users' searching and the keywords, and the keywords will be the output node of the first network model. While the second network model demonstrates a relationship between the keywords and the textual information. Each keyword is used as the input node of the second network model to form the final textual information through the conversion of the intermediate node.

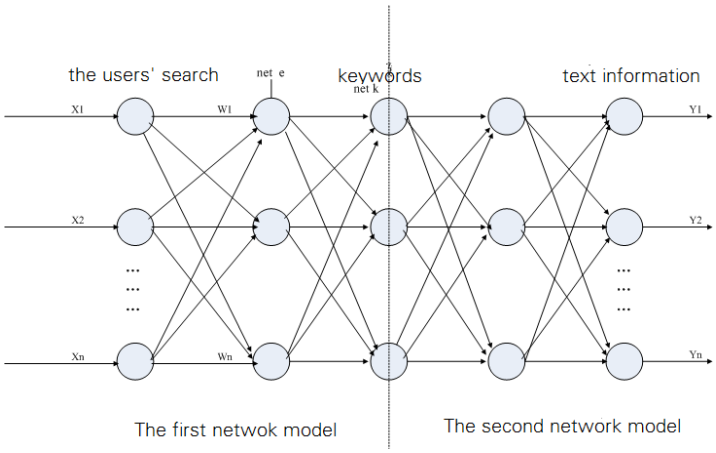

Fig.3 Multi-layer neural network model with three hidden layers

In the multi-layer neural network, we take the network model 1 as an example, in which the activation function of each hidden layer node is

$$\mathrm{net}_{ij}(t) = f\left(\sum_{i=1}^{n} w_i(t) x_i - \theta\right) \quad (i, j = 1, 2, \ldots, n) \tag{1}$$

When the information enters the first network model, the output will be given through the hidden layer to the output layer, in addition, every node requires to go through the conversion characteristic function. The S-type one of transformation functions is often selected,

$$f(x) = \frac{1}{1+e^{-x}} \tag{2}$$

The transformation function for each node in the first network model is

$$f(\mathrm{net}_{ij}) = \frac{1}{1+e^{-\mathrm{net}_{ij}}} \tag{3}$$

For the second network model, the training mode is consistent with the first one, but the input node becomes the keyword and the output layer changes into the textual information. This experiment is made to get comparisons, in which the correct rate of outputting is under comparisons in the textual information documents of different sizes among the neural networks with 1-5 hidden layers. Comparing with the accuracy of the query, the experimental results with high correct rate will be selected after this experiment. The experimental results are shown in Table 2.1, we can draw the relevant conclusions from the results of these experiments after analyzing, then select the appropriate number of hidden layers. The accuracy rate with the number of hidden layer changes are shown in Fig .4.

Table.1 Deep Neural Network model Experimental results (accuracy %)

Number	One hidden layer	Two hidden layers	Three hidden layers	Four hidden layers	Five hidden layeres
Test1 (800words)	53.5	66.7	61.2	59.0	58.2
Test2 (900words)	68.0	78.1	73.3	70.1	68.5
Test3 (1000words)	72.1	79.2	76.4	74.8	73.6

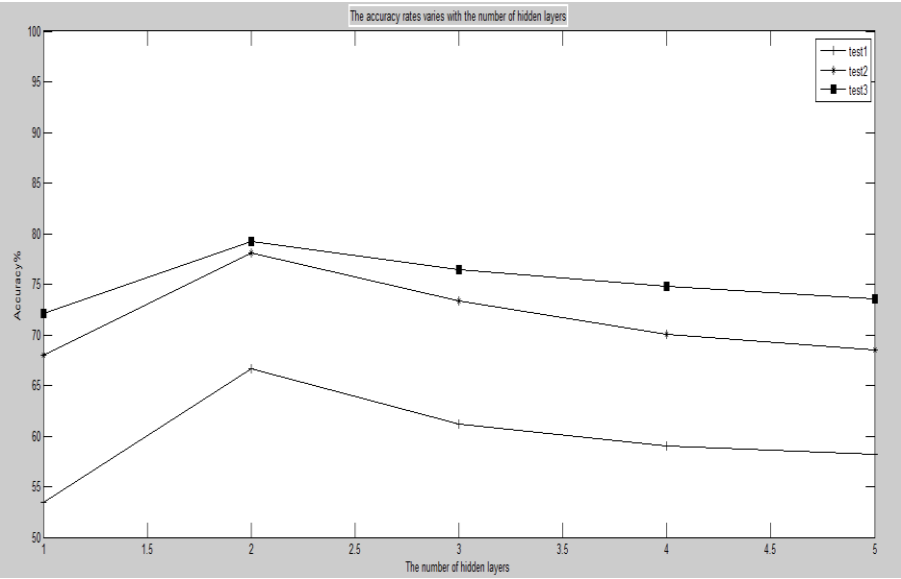

Fig.4 The accuracy rate with the number of hidden layer changes

From the experimental results with vertical and horizontal comparisons, we can obtain that it have no more obvious effects when it have more layers as for the deep learning network. The experimental results show that when the number of hidden layers is two, the search rate is the highest, and the effect is much better. So we will build a deep learning network model with two hidden layers, that is to say, a hidden layer will be added between the input and keyword layer and there is no additional hidden layer, in this way, the keyword nodes directly output through the transformation function, which can achieve best results. As is shown in Fig.5, we can get the deep neural network model in need.

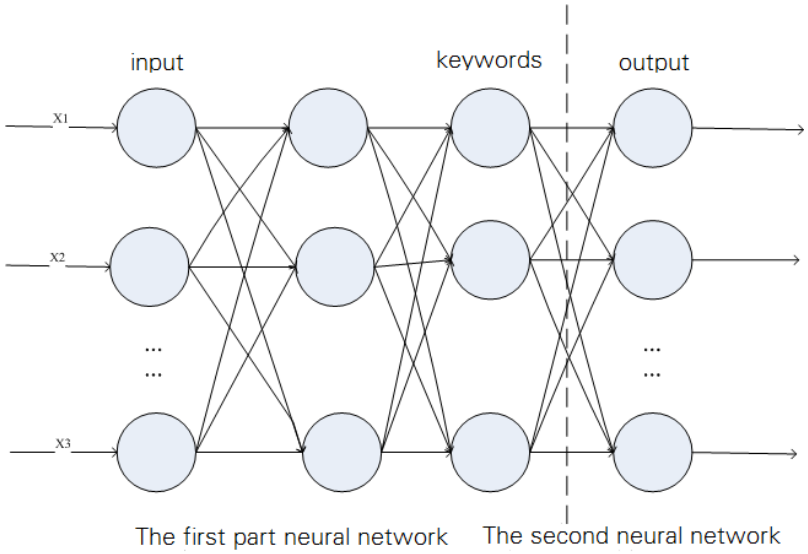

Fig.5 Multi-layer neural network model based on B-P neural network

2.4.2 The Construction of the Deep Neural Network Model Based on ReLU

There are still many shortcomings to be solved in the B-P neural network model, such as the slow convergence rate, the local optimization and the selection of the hidden point number, etc. In order to solve the shortcomings mentioned above, we can use ReLU (linear rectifier unit) to improve it. In this experiment, the multi-layer neural network model based on B-P neural network is achieved, and the nonlinear transformation function of the intermediate hidden layer node is replaced by the the ReLU linear rectification function [11]. Then during the next training process of the network model, there will be much textual information of the input terminal, in addition, the output value as zero in the middle hidden layer may also exist in this process. To solve this case, the ReLU function can be adapted to delete the output value of 0, and the associated connection can be also ignored. It not only helps the overall deep neural network to avoid the output value of 0 with output node interference,

but also reduces the number of hidden in the original model, so as to enhance the deep overall convergence rate of neural networks. As for the new neural network we have got, we can consider it as a linear classifier, therefore, the training of the classifier can be seen as optimizing the parameters of the middle to achieve the effects that the correct rate of overall can be improved. The ask of the experimental is to solve the equipment failure detection, that is to say, for the textual information, we have classified the results of the word segmentation in the previous textual processing, and we set the keywords as the K classes in the ReLU deep neural network mode, and recorded it as $C = \{c_k | k = 1,2,3 \dots, K\}$. We select N searches as training samples, give the training of the N search samples for classifier and define the training sample as $X = \{x_i \in \mathrm{L}^D | i = 1,2,3, \dots, N\}$, and the corresponding label is defined as $L = \{l_i | i = 1,2,3, \dots, N\}$. The purpose of constructing the classifier by using the ReLU function instead of the S-type function in the BP model is to improve the recognition accuracy of the users' search on the basis of the deep neural network model by optimizing the parameters in the classifier.

In next experiment, we will select the deep neural network with two hidden layers which are determined as best results in the previous experiment, in other words, the experiment will select 2-layers ReLU deep neural network. For the ith training sample, the jth neuron of the mth($1 \leq \mathrm{m} \leq 4$)hidden layer of the neural network is transformed by the ReLU nonlinear transformation function, and the resulting output is as follows:

$$y_i = \max\left(0,\ z_i^{(\mathrm{m},\ \mathrm{j})}\right) \qquad 1 \leq m \leq 2 \tag{4}$$

According to the equation mentioned above, we can see that in the forward training process, the deep neural network model will have some nodes of the output which set to 0, for the 0 points can be directly ignored in the forward training process of the entire network model, and it will have no more impacts. Then in next experiment, for the ith training samples the nodes in the entire network will be deleted as 0 and the connection with these nodes which are directly ignored, then the forward communication training [10] path of the entire network can be determined, as is shown in Fig.6 and 7. This approach will not affect the experimental results, namely, the results of the output layer of the experiment will not be changed. It will not only improve the neural network training time of the original depth, but also enhance the convergence rate, besides, the number of hidden nodes can be determined after the determination of the propagation path.

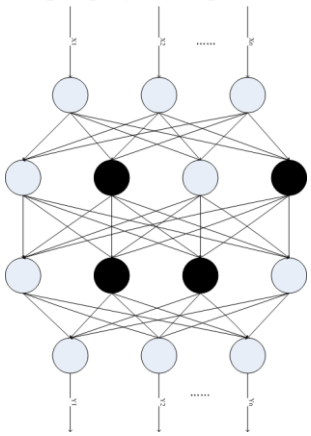

Fig.6 The propagation path of the ReLU deep neural feed-forward network

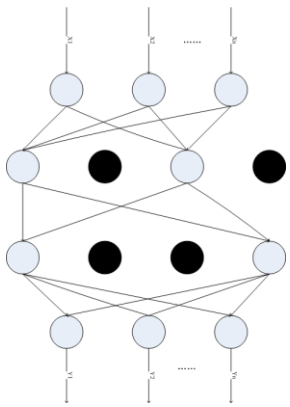

Fig.7 The final propagation path of the ReLU deep feed-forward neural network

For the neural network in Fig .7, the input of the mth layer in the ith training sample is

$$\vec{z}_i^{(m)} = \hat{W}_i^{(m)} \vec{z}_i^{(m-1)} + \vec{\Delta}_i^{(m)}, \quad 1 \le m \le 2 \tag{5}$$

$\vec{z}_i^{(0)} = z_i^{(0)} x_i$, $\vec{z}_i^{(m)}$ represents a set of all nonzero $\vec{z}_i^{(m,j)}$ vectors in the mth layer, and $\hat{W}_i^{(m)}$ represents the sub-weight matrix of all non-zero nodes in the mth layer and the open parent connected to it, $\vec{\Delta}_i^{(m)}$ represents a set of sub-paranoid vectors $\vec{z}_i^{(m,j)}$ which are non-zero, and so on, we can conclude that the output value of the last layer of the neural network is:

$$\begin{aligned} z_i^{(M)} &= \hat{W}_i^{(m)} \vec{z}_i^{(m-1)} + \vec{\Delta}_i^{(m)} \\ &= \hat{W}_i^{(m)} (\hat{W}_i^{(m-1)} \vec{z}_i^{(m-2)} + \vec{\Delta}_i^{(m-1)} + \vec{\Delta}_i^{(m)} \quad (6) \\ &= \prod_{m=1}^{M} \hat{W}_i^{(m)} x_i + \sum_{m=1}^{M-1} \left(\prod_{n=m+1}^{M} \hat{W}_i^{(n)} \right) \vec{\Delta}_i^{(m)} + \Delta^M \end{aligned} \tag{6}$$

To simplify the formula calculation, we will replace parts of the expression ways of the formula as the followings, given

$$W_i = \prod_{m=1}^{M} \hat{W}_i^{(m)} \tag{7}$$

$$\Delta_i = \sum_{m=1}^{M-1} \left(\prod_{n=m+1}^{M} \hat{W}_i^{(n)} \right) \vec{\Delta}_i^{(m)} + \Delta^M \tag{8}$$

Then we will rewritten the ReLU deep neural network of the last layer of the input formula into the following linear function form:

$$z_i^{(L)} = W_i x_i + \Delta_i \tag{9}$$

In the above formula, because the deep neural network model we used has been identified, which means the deep neural network model that contains two hidden layers. So in the former formula, M value of 2, m value of 1 or 2, while the initiative input value will be as the set of projection vectors for the keywords.

In the ReLU-based neural network model, we have divided the keywords into K classes. We will use the softmax function as the activation function of the last layer in the network. So the final output of the whole network is:

$$y_i = \mathrm{softmax}\left(z_i^{(M)}\right) = \frac{\exp\left(z_i^{(M)}\right)}{\sum_{k=1}^{K} \exp\left(z_i^{(M,k)}\right)} \tag{10}$$

y_i in the k-dimensional represents the probability that the input sample x_i is classified as c_k in the network. The physical meaning of y_i in the k-dimensiona will be expressed as:

$$p(c_k|x_i) = p(\tau(x_i) = c_k|x_i) = y_i^{(k)} \tag{11}$$

$\tau(*)$represents the sorting operation, $y_i^{(k)}$ represents the y_i of kth dimension.

The text which is used in this experiment is the same as that in the experiment of the multi-layer neural network. The main method is to increase the experimental data, to exclude the impacts of the experimental environment and make the experimental results more representative. On basis of this condition, three different experimental texts with 500, 600 and 700 words were added. These texts have different expression ways, different sizes and different generalization abilities. The experimental results are shown in the following table:

Table.2 Comparisons of experimental results based on ReLU deep neural network and multi-layer deep neural network

		ReLU-based neural network	Multi-layer neural network
Experiment number	Text size	Accuracy(%)	Accuracy(%)
1	500	66.4	65.9
2	600	66.9	66.1
3	700	71.3	74.0
4	800	66.2	66.7
5	900	80.2	78.1
6	1000	81.3	79.2

From the above table and through the horizontal and vertical contrasting, we can know that under the same textual size, the deep neural network based on ReLU improves the accuracy to a certain extent, at the same time, we can see that with the change of the textual size in the experimental environments of the ReLU-based deep neural network, the accuracy rate will also change with the increase of the textual size, besides, the accuracy rate will be also improved. As shown in Fig.8, on the other hand, in the course of the experiment the textual processing time of the deep neural network based on ReLU is significantly lower than that of the multi-layer deep neural network with the increase of the document. There are some special cases in the middle parts, but the whole time shows a decreasing trend with the increase in the presentation, as shown in Fig.9. In training feed-forward transmission neurons, the increasing trend of the whole training time is more obvious after 700 words in the documents, while the training time is steep when the document size in the between 700 words to 800 words, and the overall training time tending to rise stably within 700 words in documents, as shown in Fig.10. It can be seen that there are some shortcomings about the training time, which should be studied in the future.

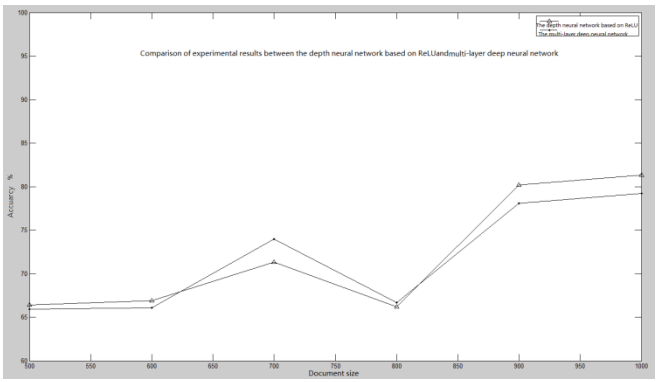

Fig.8 Comparison chart of experimental results

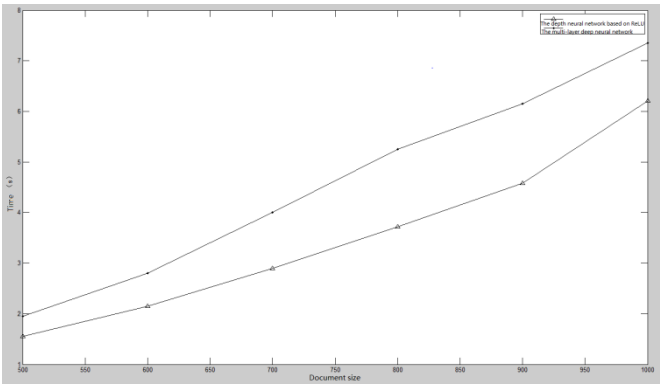

Fig.9 Comparison chart of searching time

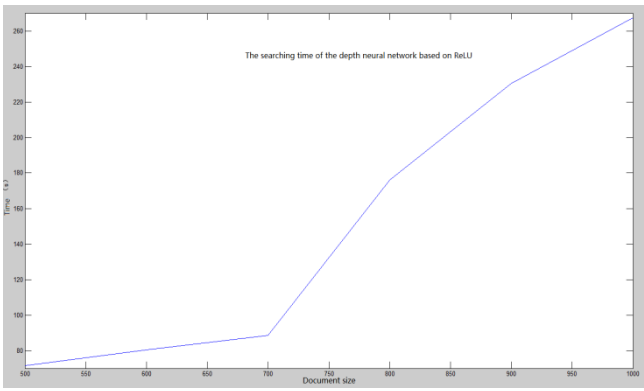

Fig.10 Forward communication training time

2.5 Optimization Ideas of Retrieval Scheme Based on Deep Learning Model

Compared with the traditional information retrieval model, the deep neural network has unparalleled advantages especially in the arrival of large data age, and the noise in the information is also increasing all the time, so the amount of information is in the process of rapid expansion. Even in the special environment of the army, with the development of information construction, the related information on the textual information about equipment is also rapidly expanding, so the document information retrieval is very difficult to get. For the deep neural network model, its own network model structure is complex and its complexity of the network model directly affects whether the model training is difficult as well as the size of the cost, so the process of the deep neural network model training is often more complex, and the cost will be relatively more expensive. Compared with the deep neural network model, the traditional information retrieval model has some shortcomings in the predicting functions including a lot of shortcomings in the accuracy rate. But the traditional model structure of the information retrieval is relatively simple, and the training time is shorter with a better efficiency. Secondly, the biggest advantage of the traditional retrieval model lies in a fact that it can process in a parallel way, which can not be attained in the deep neural network, besides, it can also solve the fatal flaws in dealing with large data in the deep neural network model. This advantage can promote the traditional information retrieval model to store the data in multiple servers, allow users to distribute the results on multiple computers without having to fuse the data, which greatly reduces the training time. Therefore, based on the present situation and the disadvantages of the former two kinds of retrieval models and based on the reality of the equipment field in the military, this paper will adopt the optimization scheme by combining the deep learning model with the traditional model, using the deep neural network retrieval model to improve the retrieval accuracy, the recall rate and the predictive function. Meanwhile, it can ensure the efficiency of information retrieval and the speed of classification of textual information. The process of combining the deep neural network retrieval model with the traditional retrieval model is shown in Fig.11:

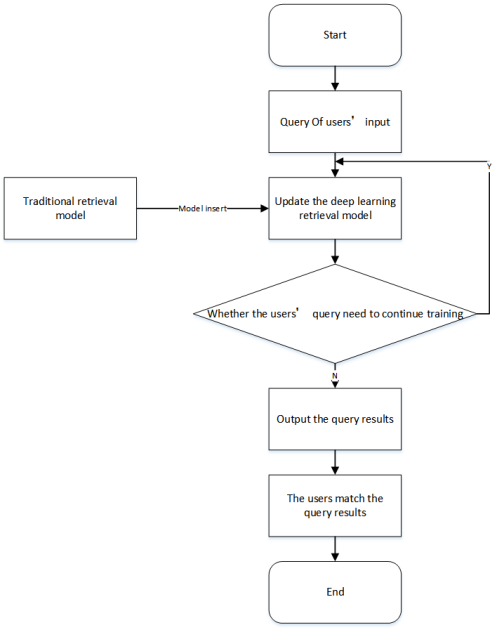

Fig.11 The process of model combination

In the process of model fusion, the user's input has been converted into the vector forms by using the relevant software, and the text of the equipment information is also transformed into the vectors. The user's input and equipment text information which is transformed into the form of vectors, can be used as the attributes of the nodes which belongs to the input layer in the model and the output layer in a deep neural network retrieval model. In the process of combining the traditional retrieval model with the deep neural network retrieval model, the traditional model is used as the model interpolation to complete the operation of the whole model fusion.

For the optimization ideas of this retrieval model, many advantages including the high accuracy rate of the deep learning retrieval model, the high searching rate and better forecasting function are utilized. At the same time, the high efficiency and fast classification of the traditional retrieval model are also used. As a model interpolation, the traditional retrieval model fuse the two models together, and the fusion process is not particularly complicated. Therefore, the idea of the optimization program show a certain feasibility.

3. Conclusion

This paper combines the characteristics of equipment information in the military equipment field, and the data mining of the textual information in the equipment inspection report is added on the basis of the original predictions of equipment failures. Because most of the predictions of equipment failures in our military equipment are based on the structured data and semi-structured data, the equipment textual information is only a small part of the large amount of information data, so it is often neglected. In this paper, the equipment information mining is added to the equipment fault predictions to further improve the prediction system for equipment faults to deal with this condition. On the other hand, based on the data mining of the textual information, this paper combines the deep neural network, which is the most popular in the field of artificial intelligence, besides, it determines the deep neural network model which is suitable for the textual information searching and the most suitable number of hidden lays. Secondly, the ReLU linear rectifier function is used in the deep neural network to replace the original transform function through the experiences. The neurons in the deep neural network model are trained to shorten the overall learning time, and the accuracy of the search results is also improved, to some extent the equipment information retrieval is optimized, according to which this paper puts forward the ideas of optimizing the scheme design which can be applied to our military equipment information retrieval. However, there are still some existing shortcomings in the design of ReLU-based neural network. On basis of the ReLU deep neural network for forward propagation, the training time still has large spans, and the training time is not stable enough. At the same time, this

paper only provides an idea in terms of retrieval optimization, and it does not propose the specific methods of model fusion. In the future study process, the learning and understanding in this aspect should be focused on in a further way.

References

[1] Seide F, Fu H, Droppo J, et al. 1-bit stochastic gradient descent and its application to data-parallel distributed training of speech DNNs[C]. Proceedings of INTERSPEECH, 2014.1058-1062.
[2] Mohamed A r,Dahl G E,Hinton G.Acoustic modeling using deep belief networks[J].IEEE Transactions on Audio,Speech,and Language Processing,2012,20(1):14-22.
[3] Dahl G E,Yu D,Deng L, et al. Context-dependent pre-trained deep neural networks for large-vocabulary speech recognition[J].IEEE Transactions on Audio, Speech, and Language Processing, 2012, 20(1):30-42
[4] Xu Kaiying, Wang Shuo, Zhang Lei, et al.A personalized retrieval model based on artificial neural network [J].Library and Information Work,2011,55(2):59-63.
[5] Zhang Junlin. Large data on the record (architecture and algorithm) [M]. Electronic Industry Press, 2014.
[6] Nan Yuhua. Network Center War concept, architecture and related research [D]. Xi'an University of Electronic Science and Technology doctoral thesis, 2011.
[7] Zhang Xingzhou.Application of Data Mining in Equipment Support Information Management [D]. Master's Thesis, China University of Petroleum, 2010.
[8] Wang Xiangbo.Study on equipment maintenance information management based on data mining [D], Ordnance Engineering College, 2013.
[9] Luo Songlin, Zhou Ping. Preliminary Study on Equipment Information Resource Management [J], Defense Technology Foundation, 2003 (5): 39-41.
[10] Qu Jianling, Du Chenfei, Di Yazhou, et al. Research and Prospect of Deep Automatic Encoder [J], Computer and Modernization, 2014 (8): 128-134.
[11] Shang Zhijun, Xue Lei, Xu Yangming, et al.Study on the study of depth learning [J] .Application Research of Computers, 2012, 29(8):2806-2810.