

The Principle and Application of Deep Learning Algorithm

Pengfei Hu^{1, a)}, Hao Tang^{2, b)}

¹ Faculty of Computer, Guangdong University of Technology, Guangzhou, 510006, China.

² Synergy Innovation Institute of GDUT, Heyuan, 517000, China.

^{a)} Corresponding author: 1114306517@qq.com

^{b)} 122260851@qq.com

Abstract. Deep learning is emerging as a multilayer neural network dimensionality reduction algorithm, the neural network model with multiple deep hidden layer through the formation of the high dimensional data of input layer feature extraction, to find a low dimensional nested data structure, the formation of more abstract effective executives said. Starting from the basic principle of deep learning algorithm, introduces the basic structure of single-layer network restricted Boltzmann deep learning algorithm and its training process. Finally, an example is given to illustrate the performance improvement of deep learning technology applied to handwritten digit recognition, and the deep learning technology is briefly summarized.

Key words: deep learning, Restricted Boltzmann Machines, automatic coding machine.

INTRODUCTION

Many studies have shown that in order to learn complex functions that express high-order abstract concepts, we need to introduce deep learning to solve artificial intelligence related tasks such as target recognition, speech perception and language understanding. Deep learning architecture is composed of multi nonlinear operation unit, the output of each lower layer as higher input from large amount of input data to learn effective feature representation, high order learning to express contains many structural information of the input data, is a good method of extraction from data, can be used for etc. the specific problems of information retrieval and classification and regression.

Deep learning is a new field of machine learning research. Based on neural network, a model of human brain analysis and learning is established, which simulates human brain mechanism to identify targets and perceive information. Its representative achievement is an automatic coding machine proposed by Hinton in 2006, which has made amazing progress in handwritten digital recognition [1]. Some previous neural network studies have also tried to solve this problem, but there are no significant results due to the lack of effective deep network models and training algorithms.

THE MODEL OF DEEP LEARNING AND ITS TRAINING

The Boltzmann machine is a kind of stochastic neural network [2], which consists of Hinton and Sejnowski in 1986 proposed a neural network based on random statistical mechanics, including a display layer and the hidden layer of the BM model. In this network, the output of the random neurons is only two states (not activated / activated). Generally speaking, 0 and 1 are used to indicate which value of the state value is determined according to the law of probability. The inherent properties of the network make it powerful and unsupervised learning ability and can learn a lot of complex rules in a large number of data.

Smolensky introduced a restricted Boltzmann machine RBM [3]. RBM is similar to an undirected graph model. The h is a hidden layer (hidden), and the v represents the visible layer. The two layers have a random number of neurons, and the state of the random neurons in the different layers is different, and the state of the different random neurons on the same layer is different, the w is the connection weight between the two layers. RBM in neurons and

neurons can be visible for arbitrary exponential distribution. We use all of the visible and hidden neurons to be two value variables. For any $i, j, v_i \in \{0, 1\}, h_j \in \{0, 1\}$.

Vector v and h represent the state of visible and hidden neurons, respectively. It is assumed that there are n visible neurons and hidden neurons in a RBM, and v_i represents the state of the i visible mental element, and h_j represents the state of the j neuron. The energy of RBM is defined as:

$$E(v, h | \theta) = -\sum_{i=1}^n a_i v_i - \sum_{j=1}^m b_j h_j - \sum_{i=1}^n \sum_{j=1}^m v_i w_{ij} h_j \quad (1)$$

The upper a_i represents the bias of the visible neuron i , and b_j indicates the bias of j in the hidden layer neuron, $\theta = \{w_{ij}, a_i, b_j\}$ is the parameter of the RBM. When the parameters are determined, we can get the joint probability distribution (v, h) based on the formula (1):

$$P(v, h | \theta) = \frac{e^{-E(v, h | \theta)}}{Z(\theta)}, Z(\theta) = \sum_{v, h} e^{-E(v, h | \theta)}, \quad (2)$$

In which $Z(\theta)$ is a distribution function.

The distribution of v in RBM $P(v | \theta)$ is the marginal distribution of the joint probability distribution of the observed data, $P(v, h | \theta)$:

$$P(v | \theta) = \frac{1}{Z(\theta)} \sum_h e^{-E(v, h | \theta)} \quad (3)$$

To determine the distribution, $Z(\theta)$ is calculated, but the operation of 2^{n+m} times. RBM is a special structure with no connection between layers. It is known that when the visible neuron state is given, the activation state of all hidden neurons is conditional independent. The activation probability of the j implicit neuron is:

$$P(h_j = 1 | v, \theta) = \sigma(b_j + \sum_i v_i w_{ij}) \quad (4)$$

$\sigma(x) = \frac{1}{1 + \exp(-x)}$ is the sigmoid activation function [4] According to the symmetry of the RBM structure, the activation probability of the j visible neuron can be obtained when the state of the hidden neuron is given.

$$P(v_i = 1 | h, \theta) = \sigma(a_i + \sum_j w_{ij} h_j) \quad (5)$$

HANDWRITTEN DIGIT RECOGNITION

Automatic coding machine is the most successful example of deep learning, which is cascaded by an arbitrary number of RBM. Its working principle is that input data is multi-layer RBM, which makes high-dimensional data become low dimensional data, completes the encoding process, generates the middle code level, and then starts from the code level to reverse the various parameters generated by the coding process, and reconstructs the input data to complete the decoding work. The whole network fine-tuning the parameters of the model by minimizing the cross entropy of the input data and reconstructing the data, making the intermediate code layer output more essential features.

The 5-layer automatic coding machine is used to classify the handwritten numerals in the data set MNIST. The process is as follows:

First, we preprocess the input data, transform the pixel value of the original digital image $28 * 28$ into $1 * 784$ row vectors, so we can build a four level automatic coding network with 784 layers in the training process while the hidden layer is 1000-500-250-30-10. Secondly, the value between 0 and 255 of the original pixel intensity is converted to the gray value between 0 and 1. Finally, 60000 training samples in the database are divided into 600 small batches of 100

groups, each training cycle processes 600 small batches in turn, and the weights are updated after the end of each batch.

The reconfiguration effect of the input number through the automatic coding machine is shown in Figure 1. The upper layer is the 15 handwritten numerals entered during the training process, and the lower layer is the handwritten number reconstructed by the automatic coding machine. It can be seen from the figure that the refactoring value is well restored to the original image.



FIGURE 1. Comparison between the original input image and the reconstructed image of an automatic coding machine

After the completion of the automatic coding machine training, we can use the labeled sample to test its classification performance and statistical error rate. Table 1 is a comparison of the error rate between the automatic coding machine and the other classifiers introduced here. It can be seen that the classification performance can be improved by using an automatic coding machine. Note that the deep learning network hidden layer and hidden layer unit number equal value will affect the classification results, and these values there is no objective inference method, therefore more flexible.

TABLE1. recognition performance of MNIST database by different classifiers

Learning Algorithm	Error Rate (%)
Automatic Coding (1000→500→250→30→10)	1.2
The BP Algorithm(500→150→10)	1.53
Nearest Neighbor Method	2.8
Linear Classifier	8.4
40PCA + Particle Algorithm	3.3
1000RBF + Linear Classifier	3.6
Support Vector Machine	1.4

CONCLUSION

Deep learning studies how to use multi-layer structures to automatically learn and represent the potential distribution features of model data. The undirected training process automatically extracts the hierarchical features required for classification and has many attributes different from the existing machine learning techniques. At present, this technology is still in the early stage of development, and some key problems remain to be solved. But through the results of automatic coding, we can see that the idea of deep learning is novel and has great potential. It is likely to have a great impact on machine learning and AI technology in the coming years.

ACKNOWLEDGMENTS

The successful completion of this paper is inseparable from my advisor, Professor Zhang Mei, who has given me many useful guidance and help in his usual scientific research and life. in the course of writing and implementation of the paper, many valuable suggestions were put forward, and her rigorous academic attitude made me benefit all the life.

REFERENCES

1. Hinton G E,Salakhutdinov R R.Reducing the dimensionality of data with neural networks. Science, 2006,313:504.
2. Hinton G E, Sejnowski T J. Learning and relearning in Boltzmann machines. Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Cambridge, USA,1986.

3. Smolensky P. Information processing in dynamical systems: foundations of harmony theory. Parallel distributed Processing: Explorations in the Microstructure of Cognition, Cambridge, USA, 1986
4. HINTON G E. Reducing the dimensionality of data with neural networks. Science, 2006, 313(5786):504 -511