Application of the MIL method and its neural network architecture

Qianruo Zhang February 2nd, 2022 897 words

Executive Summary

The multiple instance learning algorithm is understood and attempted to be coded and extended to the characteristics and development of neural network algorithms. Based on this, the MIL algorithm and neural network structure are combined to attempt regression prediction on MNIST datasets using multiple instance neural networks.

1.0 Introduction

Multi-instance learning (MIL) is typically a type of weakly supervised learning that was first defined in the task of drug molecule activity prediction[1], and is now widely used in target detection, semantic segmentation, scene classification, text classification, medical diagnosis, etc.

2.0 Construction of bags and instances

Multiple instance learning regards a bag as a sample, and each bag corresponds to a tag. There are multiple instances in each bag, whereas none tag at the instance level. Therefore, this report tends to reckon that multiple instance learning is a machine learning method between supervised and unsupervised learning but significantly differs from the semi-supervised learning method. In other words, the image with a particular label used to train classifiers does not have category tags now. Conversely, the bag has, which is quite different from previous frameworks. Another distinct point is the unique rule of multi-instance learning: if there is at least one instance labeled "+" by the classifier in a bag, the bag is positive; If all the instances in the bag are judged as "-" by the classifier, we treat the bag negative.

First, assuming that each image has a special number (id), we divide the MNIST training set into two lists based on whether the image has a label of 0 or 7, the first list contains ids corresponding to label 0 or lable 7, while the other list contains ids that are not corresponding to these labels[5].

The second step is to construct the bag, each bag belongs to one of two classes. In the first class, the bag has at least one image labelled target (i.e. containing images handwritten as digit 0 or digit 7). In the second category, bags do not have images labeled as targets.

The third step is to do a regression based on the bag level, for the numbers 0 and 7. Since the bag-space is a non-vector space, the bag-space algorithm compares the bags A and B by defining a distance function $D(A, B)$. here the k-nearest neighbor algorithm (KNN) regression method is used.

3.0 Neural network architecture corresponding to the MIL method

In recent years, as CNN has made breakthroughs in computer vision and natural language processing, more and more researchers have combined multi-instance learning with neural networks and proposed a multi-instance learning method based on neural networks. Ramon and DeRaeDt[2] first introduced the concept of multiple instance neural networks to solve the problem by using neural networks to predict the scores of instances belonging to the positive class and then integrating the instances scores into a bag of scores belonging to the positive class through a Log-Sum-Exp pooling function. A similar approach, BP-MIP, was proposed by Zhou and Zhang[3], where only the loss function and the pooling function were different.

Yan Yongyi considered the process of multiple instance neural networks to be broken down into three steps[4].

(1) Learning the instance features through the instance transformer;

(2) Pooling the instance features into bag features using a permutation-invariant multiple instance pooling layer;

(3) Finally, classify the bags based on the bag features.

4.0 Introduction and visualisation of MNIST datasets

MNIST data set comes from the National Institute of Standards and Technology (NIST), and the database is composed of handwritten digital pictures 0 to 9 and corresponding labels. The training set, 60000 images in total, is composed of handwritten numbers from 250 different people, of which 50% are high school students while 50% are staff of the Census Bureau. Meanwhile, the 10000 test sets are the same proportion of handwritten numeral data. Each image is a 28 * 28-pixel gray handwritten numeral picture. The first stage is to visual processing of the MNIST data set. In addition, we can also draw multiple sample images of a specific number 7 to see how different these handwritten samples are.

Fig. 1 Visulization of MNIST dataset.

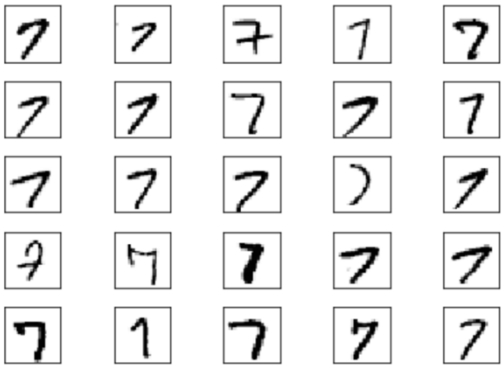

Fig. 2 Visulization of digital 7

Figure 1 shows a 2 * 5 picture, in which are pictures of single numbers 0-9 respectively.

Figure 2 shows 25 different forms of the number 7.

5.0 Recommendations

Both the implementation of the simple MIL method and the neural network structure corresponding to MIL are understood and attempted in this report, but ultimately the code is not successfully implemented to run. The report then attempted to use the KNN regression algorithm to make predictions on the MNIST dataset and the goodness of fit for this regression is 0.95, which is a more than satisfactory fit.

6.0 References

[1] Dietterich T G, Lathrop R H, Lozano-Pérez T. Solving the multiple instance problemwith axis-parallel rectangles. Artificial intelligence, 1997, 89(1-2):31-71.

[2]Ramon J, De Raedt L. Multi instance neural networks. in: Proceedings of Proceedings of the ICML-2000 workshop on attribute-value and relational learning, 2000, 53-60.

[3] Zhou Z H, Zhang M L. Neural networks for multi-instance learning. in: Proceedings of Proceedings of the International Conference on Intelligent Information Technology, Beijing, China, 2002, 455-459.

[4] 晏雍娈. 基于神经网络的多示例学习方法研究[D]. 华中科技大学, 2019.DOI:10.27157/d.cnki.ghzku.2019.001032.

[5]https://github.com/DhruvilKarani/mil_mnist/blob/acab3cf7137aa4d532f3395bef5e2e283d0653d4/bags.py#L7