La Sapienza University Of Rome

2017/2018 Academic Year

Fall Semester

Neural Networks Final Project Report

José Vicente Jaramillo

Na Zhu

Arikhan Arik

Abstract

The main objective of this project is that, creating the semi-supervised learning method for deep learning network, in the way that has been proposed in the scientific paper called **"Pseudo-label: The simple and Efficient Semi-Supervised Learning Method for Deep Neural Networks"**. This project has been created using the same techniques in the paper. Namely, Pseudo-labels, which are accepted as the true labels in the process the for picking up the maximum prediction and make the classification of the unlabeled data.[1] This method has been assumed as a prior for the semi-supervised learning technique since the algorithm uses both labelled and unlabeled data simultaneously in order to train the neural network and have more accuracy in the end of the testing phase. For the data set, the most common data set for the numbers called MNIST has been used, as the paper suggests.

1.Introduction

Neural Network types, methods and techniques have the great impact on many areas from medicine, security, to character recognition, image compression, and the usage areas of neural networks are getting widened since the performance of the algorithms are significantly better than many other methods and techniques[2]. In this project report, the main implementation of a scientific paper shall be done and the methods and techniques shall be reused in order to have similar accuracy results by training the neural network with the same methods and techniques.

The main purpose of this project is to implement the 'pseudo-labels' technique, in a neural network which has the goal of classifying the 'MNIST' dataset in 10 digits. In the later sections of this report, the methods and techniques that are been used in the implementation will be explained. Afterwards, the results that have been taken using different parameters will be revealed and compared with the source article's accuracy results. Finally, in the conclusion section, the comments and opinions about the total

procedures, the development of the project and also the results that have been taken from neural network shall be discussed.

The report will follow with the detailed explanation of the techniques and methods that have been used to develop the neural network and the explanation of the additional tools in order to achieve the development of the neural networks.

2.The Tools, Methods and Techniques

During the development of the code work for the project, Python has been used as programming language. For the tools, Numpy and Matplotlib have been used in order to declare the data structures and obtain graphical representations while Tensorflow has been used in order to apply the techniques and methods of the neural network and train it using those techniques with the dataset of MNIST.

As for the techniques in order to build up the neural network and train it for the better accuracy rate, the paper that has been followed during the development of the project, suggests several techniques such Pseudo label method, with Sigmoid unit, Rectifier Linear Unit that uses the rectifier activation function, and Cross Entropy as the loss function. For the Pseudo label a special cost function is introduced. And the alpha is a coefficient balancing between the labeled data and Pseudo labeled data

$$L = \frac{1}{n}\sum_{m=1}^{n}\sum_{i=1}^{C} L(y_i^m, f_i^m) + \alpha(t)\frac{1}{n'}\sum_{m=1}^{n'}\sum_{i=1}^{C} L(y_i'^m, f_i'^m),$$

Where alpha is defined as:

$$\alpha(t) = \begin{cases} 0 & t < T_1 \\ \frac{t-T_1}{T_2-T_1}\alpha_f & T_1 \le t < T_2 \\ \alpha_f & T_2 \le t \end{cases}$$

And T1=100, T2=600.

Every single technique that has been introduced above shall be discussed in details in the following sections.

2.1 Pseudo-labels

Pseudo-label is a method that has been used on unlabeled data for improving the accuracy of the unlabeled data prediction. The method includes semi-supervised learning algorithm, which means the integration of the labelled and unlabeled data for training the neural network. Since the neural network shall be used both the unlabeled and labelled data at the same time to train itself, the accuracy of the neural network shall have high expectation. The main idea on pseudo-label is that, training the neural network with the labelled data, then using the trained layers in order to predict the unlabeled data, therefore creating the pseudo-labels in that sense.[3] Furthermore, using the pseudo-labels and labelled data together in a new dataset to train the neural network, shall be sufficient for having higher accuracy.

2.2 Sigmoid Unit

In the paper, the sigmoid unit has been used in order to have the independence of the labels from each other, therefore the probability of each label shall be independent of the other labels.[1]

3. The Comparison of The Results

In the paper, the network has been tested with many different methods and changing the parameter of the training samples with four different sizes (100, 600, 1000, 3000). In this project, the main work has been done on the simple neural network and neural network with Pseudo-labels. The results of the pure neural network and the neural network with Pseudo-label shall be compared with the result from the scientific paper, with given different parameters for the training samples.

The comparison of the results for different training sample parameters has been shown in the following table:

-Pure Neural Network:

Training Samples	Article Results	Experimental Results
100	0.7419	0.924171
600	0.8856	0.945071
1000	0.8930	0.947575
3000	0.9396	0.947225

- Neural Network with Pseudo-labels

Training Samples	Article Results	Experimental Results
100	0.8385	0.931686
600	0.9497	0.954543
1000	0.9570	0.9591
3000	0.9720	0.9583

As observed from the two tables for pure neural network and neural network with pseudo-labels, the results of the scientific paper have been followed, and the results obtained from our experiments are quite close to each other. In pure neural network

from 100 to 1000 training samples, our experiment accuracy are higher than the paper's. And in the neural network with pseudo-labels, our experiment have better results on the 100 and 600 training samples than the paper's.
The experimental data has been taken as the average of the several trials since the dataset are randomly initialized before training.

The graphical representation of the comparison of the experimental results of pure neural network and neural network with pseudo-labels have also been presented in the following:

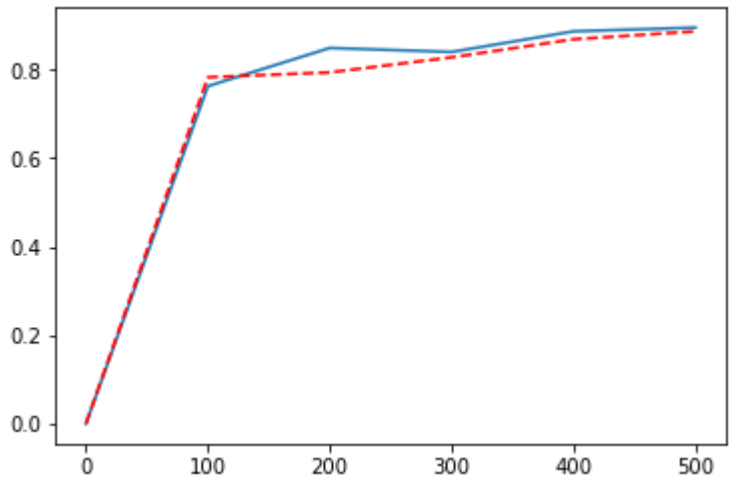


Figure 5

Figure 5 is the comparison of the pure neural network and neural network with pseudo-labels. The blue line indicates the neural network with pseudo-labels, while the red-dotted line indicates the pure neural network. This figure has been experienced for 100 training samples.

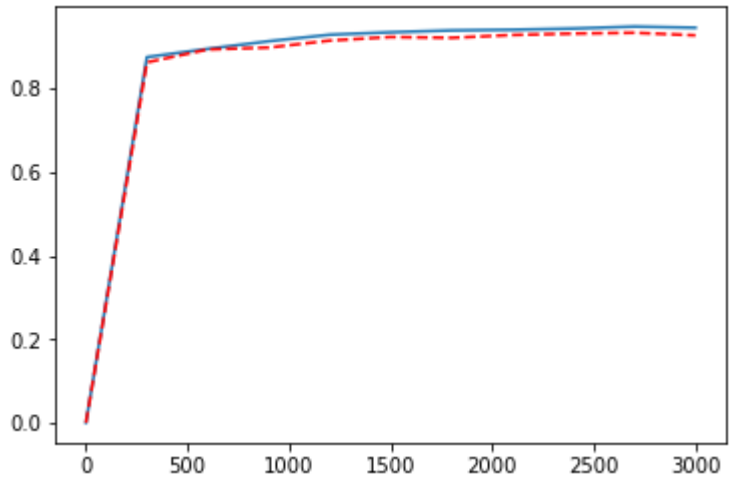


Figure 6

Figure 6 is the comparison of the pure neural network and neural network with pseudo-labels. The blue line indicates the neural network with pseudo-labels, while the red-dotted line indicates the pure neural network. This figure has been experienced for 600 training samples.

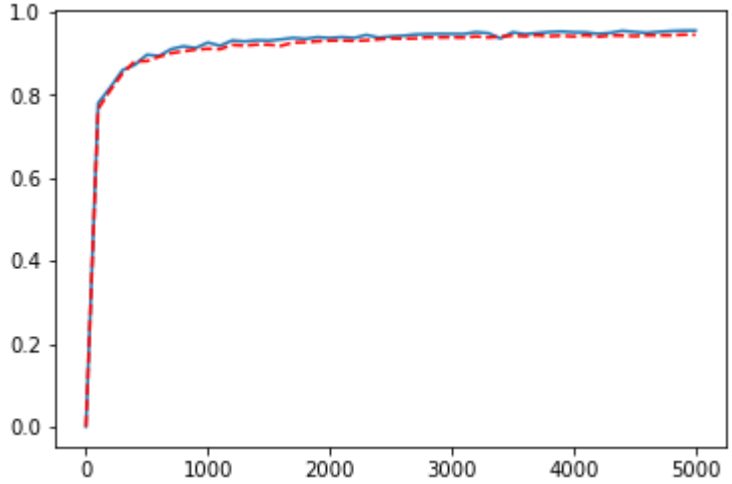


Figure 7

Figure 7 is the comparison of the pure neural network and neural network with pseudo-labels. The blue line indicates the neural network with pseudo-labels, while the red-dotted line indicates the pure neural network. This figure has been experienced for 1000 training samples.

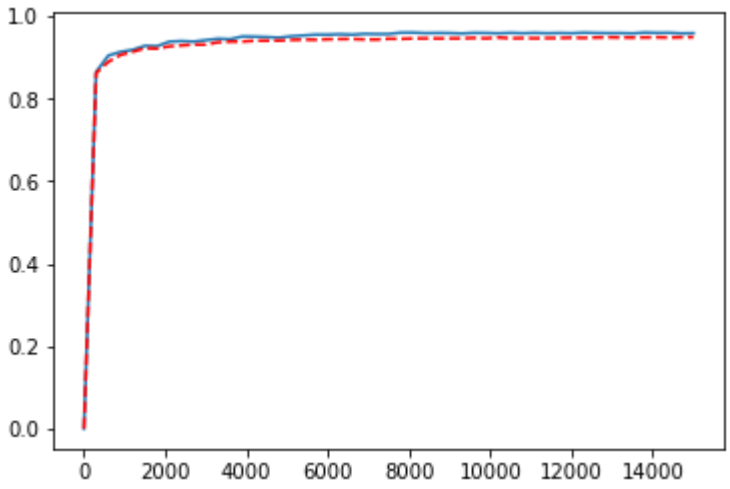
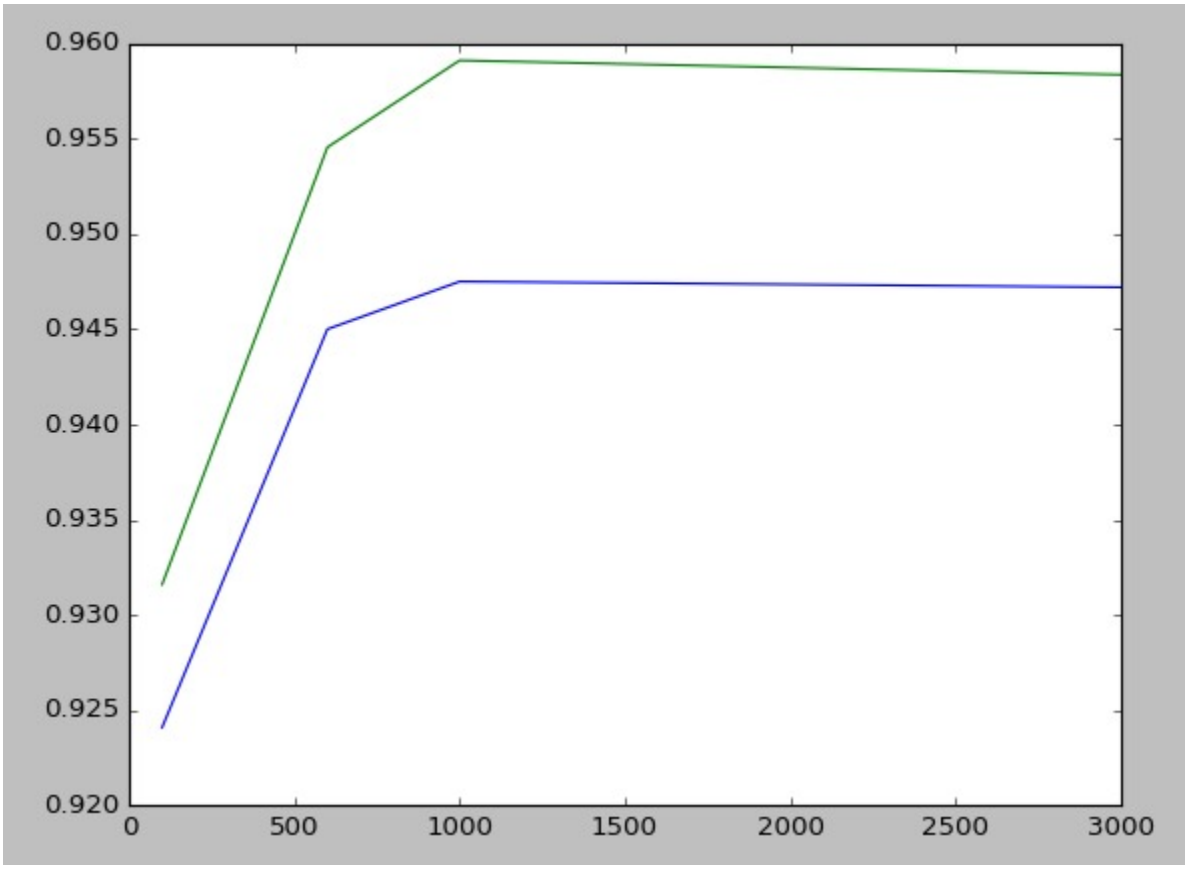


Figure 8

Figure 8 is the comparison of the pure neural network and neural network with pseudo-labels. The blue line indicates the neural network with pseudo-labels, while the red-dotted line indicates the pure neural network. This figure has been experienced for 3000 training samples.

4. Conclusion

The main objective of this project is making the same experimental work as the scientific paper suggested. The taken results are quite similar at some significant sample rate with the same neural network type, while at some points there is a big difference between our experimental results and the scientific paper data. The main difference shall be seen in the first two training samples, which our solution proposal declares much more better solution, while in the other two sample rates, the results start to get closer. The main unexpected fact in this experiment was that, the increase rate of the pure neural network and the neural network with pseudo-labels, was not increased as the paper data. However, another unexpected fact was that, the first two training sample results were extremely better than the paper results. The figure below shows a comparison between the neural network accuracy and the neural network + pseudo-label accuracy test.



Where we can see that the neural network (blue) is outperformed by the neural network plus pseudo-labels (green), for 100, 300, 1000, 3000, labeled data.

5.References

[1] Lee,D.H. Pseudo-Label: The Simple and Efficient Semi-Supervised Learning Method for Deep Neural Networks p. (1-6)
[2]https://cs.stanford.edu/people/eroberts/courses/soco/projects/neural-networks/Applications/index.html
[3]https://datawhatnow.com/pseudo-labeling-semi-supervised-learning/
[4]https://towardsdatascience.com/denoising-autoencoders-explained-dbb82467fc2
[5]https://en.wikipedia.org/wiki/Dropout_(neural_networks)