

Deep Neural Network

Philip Hoddinott

December 16, 2018

Abstract

Remove pasive voice, will, chec out rubirc The purpose of this report is to develop a neural net that can identify handwritten digets in the MNIST database at near human levels of accuracy. The neural net will be developed without the assistance of libraries such as Python's tensor flow or MATLAB's Deep Learning.

solve the MNIST on the mnist database .

The author would like to express his gratitude to Professor Hicken

for his suggestion of this project and his assistance with the methods of orbital determination through out the semester.

Contents

1	Introduction	4
1.1	Background	4
1.2	The MNIST database	4
1.3	Artificial neural network	5
1.3.1	Forward Propagation	5
1.3.2	Back propgration	5
1.3.3	activation function	5
1.3.4	Pitfalls	5
2	Implementation	5
3	Results	6
4	Conclusion	6
	Appendix	7

List of Figures

1	Sample numbers from MNIST [1].	4
---	--	---

1 Introduction

Have it solve the MNIST with a simple simple thing then try different layers and stuff

Go over tan h vs sigmoid Explain batch testing

1.1 Background

1.2 The MNIST database

The Modified National Institute of Standards and Technology database or MNIST database [2] is a database of handwritten numbers used to train image processing systems. It contains 60,000 training images and 10,000 testing images.

A number of attempts have been made to get the lowest possible error rate on this dataset. As of August 2018 the the lowest achieved so far is a error rate of 0.21% or an accuracy of 99.79%. For comparison human brains that are hardwired for pattern recognition 1.5 [Come back to \[3\]](#).

The database is comprised of images that are made up of a grid of 28x28 pixels, as seen in figure 1.

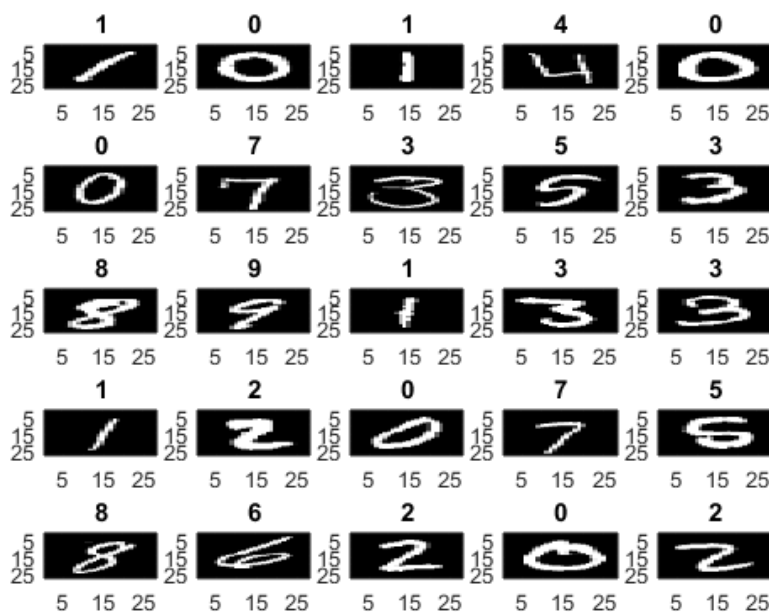


Figure 1: Sample numbers from MNIST [1].

1.3 Artificial neural network

An artificial neural network (referred to as a NN in this paper)

1.3.1 Forward Propagation

1.3.2 Back propagation

1.3.3 activation function

Tanh The derivative of tanh is seen in equation 1

$$\phi'(z) = \frac{4}{(e^{-z} + e^z)^2} \quad (1)$$

Sigmoid The sigmoid function is seen in

$$\phi(z) = \frac{1}{1 + e^{-z}} \quad (2)$$

The derivative of the sigmoid function is

$$\phi'(z) = \frac{e^{-z}}{(e^{-z} + 1)^2} \quad (3)$$

Explain importance of activation function

1.3.4 Pitfalls

The most important thing to steer clear of is over training. Overtraining occurs when the neural network trains too much to the training data. While it will have a high accuracy for the training data, its performance for the test data will decay, as it has become too well attuned to the training data.

2 Implementation

Go over how it was implemented Go over batch testing

results, comparison of different architectures

3 Results

4 Conclusion

References

- [1] Loren Shure. Artificial neural networks for beginners. <https://blogs.mathworks.com/loren/2015/08/04/artificial-neural-networks-for-beginners/>, August 2015.
- [2] Christopher J.C. Burges Yann LeCun, Corinna Cortes. The mnist database of handwritten digits. <http://yann.lecun.com/exdb/mnist/>.
- [3] John Denker Patrice Simard, Yann Le Cun. Efficient pattern recognition using a new transformation distance. <https://papers.nips.cc/paper/656-efficient-pattern-recognition-using-a-new-transformation-distance.pdf>.

Appendix 1 - MATLAB code