TJHSST Computer Systems Lab Senior Research Project: Excursions into Neural Networks 2006-2007

Tianuhi Cai

November 2, 2007

Abstract

Artificial neural networks model the circuitry of biological neurons. Inspired by the brain, they may share common properties with their biological counterparts, although they may be distinctly different. For example, the topology of the connections of neurons in an Adaline network (single-layered backpropagation network) will not be the same as those in a multi-layer backpropagation network. The size and number of layers in a neural network also influence its accuracy in its applications. In this project, different variations on neural networks will be tested and evaluated on their performance in recognizing handwritten characters.

Keywords: neural networks, computer vision

1   Introduction

Neural networks are important in that they can model highly complex, non-linear systems and can be proficient in classification and pattern recognition.

A comparison will be made of multiple types of neural networks on their performance in the recognition of handwritten characters. Handwriting samples in the form of images will be needed to train and test the neural networks. Training neural networks to recognize handwritten strokes in a video format, rather than handwritten characters in an image format, may be

done if time and resources are sufficient. Various neuron classes and neural network classes will be written in JAVA, which is suitable because of its object-oriented capabilities.

1.1   Background

In an artificial neural network, a group of simple neurons are used to display a more complex behavior as a group, which is determined by their connections and parameters. There is generally an input layer, an output layer, and possible hidden layers. The input layer takes in data, the output layer spits out data, and hidden layers do intermediate processing. Neurons fire with a 1 or a 0; when a neuron receives input, it weights the inputs by neurons connected to it, and determines whether or not to fire by the sum of the weighted inputs. A backpropagation network, in short, starts out with a set of randomized weights, and adjusts the weights on each layer depending on the error produced. In this project, the neural networks will use supervised learning. This is where the neural network is trained with a set of example pairs, and where the goal is to find a set of weights that can match the pairs.

Research on neural networks has been in existence for several decades. In particular, the use of neural networks for classification has been used.

Handwritten Digit Recognition with a Back-Propagation Network by Y. Le Cun et al at AT and T Bell Laboratories used a large back-propagation network to read hand-written zip codes. It took inputs in the form of black and white images. The project focused on the effect of architecture on the ability of the network to recognize digits. Its architecture was modeled in such a way that not all neurons from each layer were connected; there were five layers in total, with some neurons receiving only local input. This architecture performed well in identifying the handwritten digits, regardless of position.

1.2   Expected results

During the course of the research project, variations of the neural network will be created and tested. It is expected that whereas a simple, one-layer neural network will not be able to handle noise or variation in the incoming data, a more advanced neural network will perform better.

The goal of this project is to learn about different types of neural networks and their strengths and weaknesses. It is also to gather evidence of which

type of neural network is the most effective at handwriting analysis.

1.3   Type of research

Use-inspired basic research, to pursue fundamental understanding but motivated by a question of use (Pasteur's work on biologic bases of fermentation and disease)

2   Procedures and Methodology

Over the course of the next few months, a basic neural network will first be built. Then, more advanced and complicated neural networks will be created from that. Research, design, programming, and testing will all be done throughout the year as different networks will be created. Handwriting data will be needed for testing and training neural networks. For the language, JAVA will be used, and Eclipse or JGrasp would be used for software.

The project will be deemed successful if the neural networks created can successfully identify handwritten characters.

To test if the program works, a neural network will be trained with test data. Then, the network will be tested with data that was not from the training sample. If the program can accurately identify the characters, it will be successful.

To aid in the portrayal of the data, the weights of the connections between neurons can be continuously printed out, and then graphed, possibly using GNUplot.

To test the performance of a neural network, it will first be trained with a set of images. Then, it will be run to identify another set of test images. If it identifies the test images correctly, it will have performed correctly.

The requirements for the program would be to create a functioning neural network, and preferably with one that could correctly identify handwritten characters.

3   Expected Results

It is expected that whereas a simple, one-layer backpropagation neural network will not perform well on identifying characters, a more complex neural network may perform better.

An exploration in the different types of neural networks may help future researchers and/or students by presenting empirical data on the performance of variations on the neural network.

4 Discussion, Conclusion, Recommendations

The purpose of this project was to compare different neural network structures in their performance in recognizing handwritten characters.

5 Appendices

6 Literature Cited

References

[1] Le Cun et al, "Handwritten Digit Recognition with a Back-Propagation Network", AT and T Bell Laboratories, 1990.

[2] Anna Hart, "Using Neural Networks for Classification Tasks - Some Experiments on Datasets and Practical Advice", Department of Mathematics and Statistics, Lancashire Polytechnic, 1992.

[3] Stuart Russell and Peter Norvig, Artificial Intelligence - A Modern Approach, Prentice Hall, 2003.

7 Acknowledgements

I had a lot of help from Mr. Latimer, who provided knowledge and guidance.