

AI Final Project: ANN Handwriting Digit Recognition

Hector Zhang - 100947935

Problem Domain

In the field of optical character recognition, Handwriting Digit Recognition is most popular problem to the area of machine learning in AI. For this project, the goal is to complete electronic conversion of images, hand written or printed text by over thousands unique authors, into machine-encoded text, whether from a scanned document, a photo of a document or from subtitle text superimposed on an image.

Depended on the scope of the field, it will be only considered a subset of characters, digits 0 to 9 in this project. This narrowed down domain can be applied to the identification of US postal codes, bank check identification processing, and any task that follows numeric data processing. In order to get formally state the problem domain, the final AI project will be a handwriting recognition program specifically targeting digits zero to nine, on top of a machine learning algorithm accomplished by artificial neural networks.

Motivation

During the semester, we've learned many techniques to employ artificial intelligent algorithms to search for finite existent (or non-existent) goals within the domain of puzzles and two player games. These types of exercise build rules around the domain and programs that solve a particular narrow task with a high rate of success. I would like to explore something new AI technique which is outside of the realms of definite rule-based game systems and deploy an artificial agent that learn in a way more like the biological, ourselves.

The main motivation is to create a project that can solve my problem without being explicitly programmed. I've had an enjoyable experience with heuristic-driven intelligence who can be only as smart as the programmer themselves. It seems to take up this project to challenge myself in creating an intelligent agent that tunes itself based on experience. This type of agent is known to simplify my problem domain when compared to the performance of ordinary, rule-based programming. I would like to leverage my interest and existing experience in the subject to pursue this topic independently (until the professor teaches it) to learn and experience machine learning and artificial neural networks first hand.

Techniques

In this section, I will go over the general design and architecture of my system and talk about the secondary sources of data in which I require for the operation of my project. This project, as mentioned above, will consist of building an artificial neural network to learn inputs based on experience.

My neural network will consist of an estimated about 5000 input units that each corresponds to a pixel of my 28x28 image and 10 output layer units which range from units 0 to 9. Right now, it is not clear to know how many hidden layers, or hidden layer nodes I will discover.

The purpose of this hidden layer is to detect groups of pixels to divide and conquer an image into deciding if it is a part of some digit. They will be programmed in such a way to return true if the group of pixels it's responsible for matches the characteristics that that hidden layer node is looking for.

I propose to implement the backpropagation algorithm to learn the sample data that I provide. To propose the details of this algorithm, to the current extent of my knowledge, back propagation has two phases in which is used to learn.

- The first phase propagates the inputs from the first layer (picture pixels) forward into the hidden layer(s) which will be derived for error and will trigger an activation function if the data pass the test (a current setting of weights and bias).
- The second stage will adjust the weights and bias based on error and output its findings into the output layer.

My program will repetitively apply this algorithm until it's able to successfully determine digits from the images that are provided. In order for this system to work, I will need to feed a series of inputs into the program for the network to learn and differentiate between input. All input has been normalized to consist of a 28x28 pixel, black and white image with a legible digit written in the center. I will be leveraging the MNIST database of handwritten digits, which consists of 70,000 handwritten samples, to train and test the outcome of my project.