1 Introduction

In many instances, it is extremely convenient to be able to approximate one function by another function. For example, a certain function may be extremely difficult or inconvenient to evaluate, but a suitable approximation may be readily evaluated without sacrificing much in terms of accuracy. In particular, a digital computer may evaluate a transcendental function by using an approximating polynomial. Another context in which function approximation arises is when a function is itself is unknown, but its output for several input values has been observed. In this case, the task of function approximation amounts to the problem of how best to infer the output of the function for input values that were not observed. A related problem involves estimating the parameters in a function when the observed values were subject to measurement noise. Finally, an approximating function may distill important information about the function. Suppose that a transmission selectively attenuates certain frequencies. By decomposing a signal into a sum of sinusoids and modulating the amplitudes of the sinusoidal components corresponding to the attenuated frequencies, it is possible to determine what signal will be transmitted across the medium. These brief examples are just some of the many reasons that function approximation is a valuable mathematical tool.

Due to the importance of function approximation, many techniques have been developed to deal with this problem in different contexts. This project will explore some of the more basic techniques commonly used for function
approximation and then apply these results to a non-trivial application of practical interest.

2 Theory

The ability of polynomials to approximate continuous functions was established by Weierstrass (1885). Because this paper is offered jointly to the Mathematics and Computer Science Departments, one aspect of the work will focus on a thorough understanding of the proof of Weierstrass’s Approximation Theorem. While Weierstrass’s Theorem is theoretically appealing, it can be applied in very limited circumstances because it requires explicit knowledge of the function to be approximated. The next step in the theoretical development presented in this paper will investigate what techniques are used when the exact form of the function is not known, but rather the values of the function are known at a number of sample points. First, the case of a linear approximation will be considered in the presentation of linear least-squares. Subsequently, the linearity constraint will be relaxed and the paper will delve into the nonlinear technique of the method of gradient descent. The investigation will culminate in the presentation of the artificial neural network, a paradigm for nonlinear computation that is often used in conjunction with gradient descent for function approximation.

3 Application

After investigating the theory underlying function approximation, the paper will proceed to consider a non-trivial application of those techniques to an interesting problem – optical character recognition. The United States Postal Service has employed artificial neural networks effectively to perform pattern recognition on handwritten zip codes in order to classify mail. This project, being much more limited in computational resources and constrained in time, will be far less ambitious, with the goal being merely to identify a comparatively few number of computer-generated characters. Nevertheless, this problem will illustrate the fundamental concepts used in more elaborate systems and will prove that the methodology is viable.
4 Deliverables

The paper component detailing the theoretical development of the approximation techniques will constitute a significant portion of the work for this project. However, the paper will also seek to develop theory and practice jointly, presenting actual simulations and approximation results for each method that is presented. Thus, the deliverables will include programs capable of performing the various types of approximation that are presented throughout the paper. The primary goal of the project is still to create a functioning optical character recognition system and this will also be the most substantial deliverable.