Algorithms and Hardware for Implementing Artificial Neural Networks

Nathan Hower

Abstract

Complex problems require sophisticated processing techniques. Artificial neural networks are based on the communication of neurons in living brains. Like the millions of neurons in your brain, these models often require a parallel processing approach to be computed at practical speeds. Artificial neural networks are being used in a growing number of research fields, and the development of algorithms and software for ease of use will lead to advancements in dozens of areas. One such area is climatology and weather prediction, and research is proposed which will establish a system for using neural networks in climate simulations.

1. Introduction

The astounding complexity of nature can serve as inspiration for technology. Artificial neural networks are computational models inspired by the intricately interconnected collection of neurons that work together in a biological brain. The importance of artificial neural networks is apparent in their versatility, as they are being used for computation and data analysis in a wide range of applications. The concept arose out of neurologists in the 1940's using electric circuits to attempt to model the activity of neurons in the brain. The development of computers a decade later created a platform to explore neural networks further. The use of artificial neural networks as a computing device was overshadowed by the now-standard von Neumann architecture, and it was only in the 1970s that research and development of neural networks resumed, allowing them to extend over all areas of science in their processing abilities.

This paper gives an overview of how a neural network functions and discusses the hardware being used to run such computations. Section 2 gives a simple example of a perceptron, which is the basis of a neural network. Section 3 details the parallel computing tactics being used to process the networks. Section 4 describes several subtopics that must be considered to increase computational power and functionality. Finally, Section 5 proposes research to establish an efficient method of using artificial neural networks in climate prediction and simulation.

2. Perceptron – The Artificial Neuron

Although artificial neural networks are based on biological systems, their use in computation can have a wide range of requirements. The type of neural network needed is determined by analyzing three main factors: inputs, transfer function and output neuron count. This information forms what is known as a perceptron. A perceptron is a mathematical model of an individual neuron, and complex neural networks are built by using the output of one perceptron as the input for the next, creating long chains of computation. An input value is commonly represented over

a numerical range. Each input node has a weight or bias associated with it. The input is multiplied by the weight, which is determined by the importance of that particular input. The results of each input node are summed and processed by a transform function. This modifies the data to comply with the required output format. An example of one such algorithm is:

1. If the summed value is greater than the numerical threshold X, return 1
2. If the summed value is equal to the numerical threshold X, return 0
3. If the summed value is less than the numerical threshold X, return -1.

Each transform function will result in one output value of either 1, 0, or -1. The next layer of nodes may sum these outputs and pass the data through another transform, returning the result. This process continues until the final output set is processed. The ways in which these nodes are connected, as well as the nature of the transfer functions used, depend on the specification of the desired output.

3. Parallel Computing Systems

The inherently parallel nature of neural networks requires hardware that is significantly different from the computers we are used to using in everyday life. The current focus on parallel processing has resulted in computer architectures that are well suited to the computational needs of artificial neural networks. Like any area, parallel processing seeks to increase the efficiency and speed at which problems are solved, but optimal solutions are often monetarily expensive creating a tradeoff between power and affordability. Balancing these two techniques can result in greater computational efficiency, although this will differ depending on the problem and type of neural network being used.

3.1 Multiprocessor Computers

One potential solution is a single computer with multiple processing units. According to Seiffert [1], these have the fastest data transmission speeds between nodes as the systems are designed with processor communication in mind; however, there are several drawbacks. Multiprocessor computers are typically very expensive to build, and contain a limited number of processors and therefore can only handle a limited number of nodes.

3.2 Heterogeneous Clusters

A collection of networked computers or processors is referred to as a cluster. A heterogeneous cluster consists of any number of connected computers that may be of vastly different types. Distributed computing systems are an example of this type of clustering as the composition of the individual computers is generally arbitrary. This nature allows for any number of nodes as

more computers can be added into the network at any time [1]. The cost of these processors is very flexible as any type of processors can be used, ranging from average personal computers to more advanced systems. The major drawback is that there may be additional overhead caused by connecting disparate types of computers

3.3 Beowulf Clusters

Another type of architecture known as a Beowulf cluster provides the best cost to performance ratio available today [1]. Each node in this type of cluster is an identical processor. This eliminates the overhead drawbacks of heterogeneous clusters while maintaining their capabilities of being indefinitely expandable and of utilizing inexpensive personal computing hardware.

4. Current Obstacles

4. 1 Available Software and System Tools

A result of the recent nature of artificial neural network development is that there is not an abundance of software and systems tools available for researchers. This increased difficulty can deter scientists from using neural networks in their work as they may prove too complicated to implement. Some research departments have dedicated a portion of their efforts into developing software that is needed or useful to their studies, and this software can often be applied with little to no modification to other projects. By releasing these programs, other researches can benefit from them and allow their work to progress without being hampered by technological restrictions not directly related to the problem. Long and Gupta [2] report on a software unit known as the Scalable Parallel Artificial Neural Network that incorporates backpropagation learning and facilitates fast node communication. It is scalable because it is designed to be applied to a variable number of nodes. LeBlanc [3] details research at the University of Rochester in virtual reality and robotic vision that led to the creation of systems tools that reduce setup time and overhead costs of node communication.

4.2 Scalability

A concern with current techniques is that they are often unadaptable to more complex situations, which often requires a modified program or computer setup to scale to problems of a larger scope. Peck [4] write about the Large Edge Node Simulator, which is a specialized environment for simulating biological neural networks. The methods used are also applicable to artificial neural networks. By analyzing particular edge nodes, workably accurate results can be obtained while eliminating some of the computation usually needed. The program's abilities for data management and processing also add to the growing collection of available software for neural network research.

5. Artificial Neural Networks for Climate Prediction/Simulations

I propose to investigate the potential of artificial neural networks for use in applications modeling climate and weather. Climate affects everyone and everything on Earth, and is the result of a complex interaction of factors. Currently, the best method for predicting possible climate effects is through simulation. Collecting observational data is the first step in this process. This information has been meticulously compiled from many locations on the earth and becomes more detailed as additional influences are considered. The sensor technology to cultivate this data is advanced and returns are massive and always growing. The next step is to extrapolate probably outcomes of future weather from this data. This requires powerful computing devices and techniques for simulation.

5.1 Why Artificial Neural Networks?

Any scope of climate analysis, from individual lake or forest ecosystems to larger patterns affecting the entire globe, relies on an immense number of factors. These variables are highly connected and slight changes in one have been shown to affect seemingly unrelated weather patterns. This makes artificial neural networks an ideal candidate for processing the observational data we currently possess and continue to gather. As new variables are discovered, they can easily be added as additional inputs to existing models. Neural networks have been used successfully for analyzing small aspects relating to climate such as the effects the presence of carbon and nitrogen in an environment [5]. Another area neural networks have proven useful is calculating probabilities associated with climate [6].

5.2 Testing – Network Types and Learning Algorithms

The main method proposed for this research is the testing of neural network variants to determine efficiency. Many network types as well as learning algorithms will be examined in as many combinations as possible. Some of the promising network types that will be used are radial basis function networks, useful for time series analysis and prediction, the self-organizing properties of Kohonen maps, and stochastic networks, which introduce random variation in calculating probabilities.

The extent of this testing is broad. The initial approach will analyze small subsets of climate variables that are closely related. The unique relations of specialized areas such as the precipitation cycle or the transfer of organic molecules from deceased organisms to the environment may be most efficiently computed by different types of networks. Once this is complete, the second stage will involve connecting these subset groups into increasingly larger climate models, resulting in hybrid networks utilizing a variety of organization types. The goal of this is to ultimately build large scale models that function similarly to existing prediction simulations and have additional capabilities unavailable to current techniques.

5.3 Implications of Results

The purpose of this research is not to directly advance the analysis of factors and interactions in weather, but to provide a powerful framework for producing more accurate simulations. One challenge will be obtaining usage of sufficiently capable computing devices for the testing of network types. One possible result is that it may be discovered a specific computer architecture must be designed to fully implement a neural network climate modeling approach. The role of neural networks in science as a useful tool for processing data in various fields appears promising in climatology.

References

[1] Udo Seiffert, "Artificial neural networks on massively parallel computer hardware," European Symposium on Artificial Neural Networks, pp. 319-330, 2002.

[2] Lyle N. Long and Ankur Gupta, "Scalable Massively Parallel Artificial Neural Networks," Journal of Aerospace Computing, Information, and Communication, Vol. 5 Num. 1, pp. 1-11, 2008.

[3] Thomas LeBlanc, Dana Ballard Christopher Brown, Randal Nelson, and Michael Scott, "Rapid prototyping of parallel robot vision systems using virtual reality systems and simulation," Progress Report and Continuation Proposal, pp. 1-5, 1998.

[4] Charles C. Peck, James Kozloski, A. Ravishankar Rao, and Guillermo A. Checchi, "Simulation infrastructure for modeling large scale neural systems," IBM T.J. Watson Research Center, 2003.

[5] Thomas A. Clair and James M. Ehrman, "Variations in discharge and dissolved organic carbon and nitrogen export from terrestrial basins with changes in climate: A neural networks approach," American Society of Limnology and Oceanography, pp. 921-927, 1996.

[6] R. Knutti, T.F. Stocker, F. Joos, G.K. Plattner, "Probabilistic climate change projections using neural networks," Climate Dynamics, Vol. 21, pp. 257-272, 2003.