IDEA GROUP PUBLISHING
701 E. Chocolate Avenue, Suite 200, Hershey PA 17033, USA
Tel: 717/533-8845; Fax 717/533-8661; URL-http://www.idea-group.com
ITP4525

Artificial Neural Networks (ANN): History and Foundations of Traditional ANN

Najib Saylani
Hofstra University, Frank G. Zarb School of Business
134 Hofstra University
Hempstead, New York 11549-1340
Phone:(516) 463-5716, Fax: (516) 463-4834
acsnzs@hofstra.edu

ABSTRACT

It all started with extreme optimism and ended with progressive respect for the system. A system, researchers were eagerly trying to fully understand and model. Today, models and algorithms that support them may have all failed to solve the main problem of completely and successfully modeling the human brain's processes. But, most of these models and algorithms have found a sanctuary in applications spanning fields from science and technology to business. In the business field, their applications span areas in Accounting in the process of auditing, in Finance they are used for Credit card approval and fraud detection, in Management their main application is in assessing corporate merger possibilities, in Marketing they are useful for customers' profiling of their purchasing patterns and the list go on.

INTRODUCTION

Researchers in the past and present have worked on different models of the brain. But in biological systems, modeling many of the brain functions that process stores and recognizes information is still facing a huge set of obstacles. The problem of consciousness is the most famous one. Scientists have at their disposition a good description of a single neuron, at least as far as its internal structure and its molecular composition. They have designed a fairly good map of the brain functions. They can keep track, record and analyze neuronal signals across the cerebral cortex, but they are still far behind in designing a model which will best mimic the human brain.

It is understood that the brain is a set of neurons densely connected as well as being a system capable of massive distributed processing. It exhibits a surprising capability of self-organization and stability by a-not-yet completely known process of self-interactions and interaction with its environment. For most researchers in this domain, the brain is a non-linear dynamic system capable of stability and equilibrium. Neurophysiology uses mathematical models to study specific functions of the brain. These models simulate the flow of a signal from one neuron to another across axons and their effect at the surface level of a given neuron. But what has not been fully developed yet is a model that simulates the brain behavior as a whole. Understanding individual parts of the system does not explain the collective processes of a collection of interconnected neurons.

Scientists from biology, psychology, computer science, mathematics and physics are all contributing to designing a working model of the brain. Together with neuroscientists, they are working on understanding how the neuronal system of the most primitive being (with few neurons) works and interacts with its environment. A single neuron in beings such as the paramecium is still a complex system. It looks as if, even a single unit (neuron) is a complex entity. Few are too much to handle. The attraction to even partially exploring and explaining the brain functioning, lead others in the fields of Computer Science, Mathematics and Physics to supply neuroscientists with the right tools needed for a fast and efficient analytical processes. At the same time they benefit from their partial results by exploring their application to a variety of problems where other approaches had previously failed.

In any case, all apparently agree that the cerebral cortex is a fast processing system capable of simultaneous parallel and distributed processing. Current results in cellular biology and physiology suggest that a neuron is more than just an on/off switch namely a system where many molecular signals are generated in a more complex structures not fully understood. Recently, the discovery of fine neuron structure called micro-tubules and the theoretical opinion that these micro-tubules are capable of some type of "photonic transportation"-via some type of signaling-encouraged many scientists to explore new ideas in dealing with the brain. The scientists are so optimistic that they started attacking the problem of consciousness and the binding problem, which is the capability of the nervous system to, not only recognize a pattern but also to generate or associate with it information or other patterns generated from other locations of the brain.

HISTORY

Since the identification of a different type of cells in biology by Ramon Y Cajal in 1911 (who called them neurons)[9] biologists recognized the need to describe and understand both the structure and functioning of the brain. It was clear to them, as far as its structure is concerned, that the brain is a very complex organ with massive interconnections between the neurons. The brain is a network of neurons. But as far as functionality, little is understood. In parallel with the work of biologists, neuro-psychologists and psychologists in general through extensive research proposed various theories through various experiments mostly in behavior and consciousness.With the help of mathematics, models were created for a single neuron to study its functions and its relation to its environment [1].

It was not until the 1940's that the interest in modeling the brain reached its peak with McCulloch and Pitt [5,9] with their seminar on neural networks. Rosenblatt [5] followed in 1960 with the Perceptron. In 1943 McCulloch and Pitt introduced a two layers Perceptron as a basis for neural modeling. During the same period, Minsky and Papert [3,5,9] with their work on logic of threshold networks realized that their network as a finite state machine was capable of universal computations.

The Perceptron of Rosenblatt introduced the idea of weights along neural connections. By choosing the appropriate weight a network can learn how to solve a particular problem. It is the Perceptron of Rosenblatt that was criticized by Minsky and Papert as not being able to solve the XOR problem (1969). After this, Rosenblatt tried to overcome the problem with a Perceptron with multiple layers, but lack of a suitable algorithm at that time threw the field of neural network into complete darkness until the 80's [2].

In 1956 Von Neumann [5,6,9] played a part in developing the field of ANN by suggesting the use of redundancy when making reliable network, with unreliable parts. When one or a certain number of parts fail the whole system should be able to continue functioning using the working parts (Redundancy in structure and function). As a consequence of these optimistic researches, Von Neumann work had contributed to the creation of computers known today as the Von Neumann machine. While Minskey and Papert contributed to freezing research in approaching the creation of intelligent machines with neural network modeling, other researchers continued their efforts in the 70's that leaded to the introduction of associative contents-addressable memory. It is important to mention that many researchers, not associated with computer science, studied the general problem of learning and contributed indirectly to a comprehensive development of ANN. All in parallel developed different approaches based on their main field of study [2].

Marr (1969-1971) [9] with a biologist approach, developed new theories of the functioning of the cerebellum and hippo corpus, and contributed to locating and assigning specific functions to specific area in the brain.

Malsburg and Cooper around 1973 [9] focused their study on the functioning of the visual system. Everyone was trying to understand the nervous system. Physics got involved when Cragg and Temperley (1954-1955) [7] first redefined and reconstructed McCulloch-Pitts network, this time, by constructing a set of interconnected neurons. Each neuron is subject to a certain magnetic orientation that defines its spin. Each neuron's spin would contribute to the spin of the whole network and to a stored pattern.

The field of ANN renaissance started in 1981 by incorporating ideas from physics, especially statistical mechanics. Hopfield [5,6,7,9] introduced an energy function and the notion of dynamic local attractors. Then between 1983 and 1986 Hinkron Sejnewski and Peretts [7] used stochastic units in their study of neural networks. Then, still borrowing ideas from statistical mechanics, Amit in 1985 [7] developed a system called spin glasses through methods from the field of random magnetic system. In 1985 Rumelhart [2,3,5,7,9] and others went back to study the Perceptron with multiple layers and introduced a new method of learning called the back propagation. The back propagation approach is a method widely used and explored in today's applications and research in ANN [2, 3].

The area of ANN is still evolving by continuously accommodating ideas from statistical mechanics. New currents of thoughts are being explored with sources in the fields of physics, particularly the field of quantum mechanics [8,10].

HISTORICAL AND CURRENT APPLICATION OF ANN

As far as applications of ANN, pattern recognition is the most famous one. For pattern recognition an ANN will be presented with many variations of the pattern and should be able later converge to the desired one in the retrieval process even if it is represented with a slightly modified one (pattern completion). Machine vision and speech recognition are also some of the main area of applications.

Currently, the applications span other fields, such as solving linear systems of equations that map parallel programs onto a hyper cube or space mapping, signal processing, where ANN are used to solve partial differential equation, filtering and classifications of numerical parameters and others.

While most of business applications of ANN is part of what we label as "emerging technologies", the serious historical applications of ANN do span at least the period overlapping the past teen years. One of the most famous problems that were tackled by the use of ANN is the Traveling Salesperson Problem (TSP). TSP is a problem of optimization [4] and ANN are good at this type. During this period, some successes of applications of ANN in the engineering areas were duplicated in business to support, and sometimes to replace, AI and other Expert System approaches to building knowledge based decision support systems. ANN are very good at pattern recognition and completion. In business, anytime a decision maker is faced with finding specific patterns, trends or making sense of incomplete or difficult to relate set of historic data, ANN can be candidate for an application. In today knowledge based decision support system, ANN are one set of tools part of an integrated system combining AI, genetic algorithms, fuzzy logic and expert systems. In Accounting in ANN can be used in the process of auditing and finding abnormal transactions, in Finance they are used for Credit card approval, fraud detection, clients' credit scoring, loan approval and market forecasting. In Management their main application is in assessing corporate merger possibilities. In Marketing they are useful for customers' profiling of their purchasing patterns and data mining. The list includes application in the area of human resources and operations. The list of application of ANN in business is large for a paper restricted to this size to cover. But, a book by Efraim Turban and Jay E. Aronson [11] does a better job in this area.

BIBLIOGRAPHY

[1] Aleksander, I.D. (1989). Neural Computing Architecture, the MIT Press.

[2] Dayhoff, J., Omidvar, O., (1998). Neural Networks and Pattern Recognition, Academic Press.

[3] Fischler, M. A., Firschein, O., (1987). Intelligence: the Eye, the Brain and the Computer, Addison Wesley.

4[13] Giarratano, Riley, (1998). Expert Systems, Principles and Programming, 3rd Ed., PWS Publishing Company.

5[4] Hayken, S., (1994). Neural Networks: A Comprehensive Foundation, Macmillan.

6[5] Hecht-Nielson, R., (1989). Neurocomputing, Addison Wesley.

7[6] Hertz, J., Krogh, A., Palmer, R. G., (1991). Introduction to the Theory of Neural Computation, Addison Wesley.

8[9] Lockwood, M., (1989). Mind, Brain and the Quantum, Blackwell.

9[12] Rumelhart, McClelland, and the PDP Group, (1988). Parallel Distributed Processing, Explorations in the Microstructure of Cognition, Volume 1 & 2, MIT Press.

[10] Stapp, H. P., (1993). Mind, Matter, and Quantum Mechanics, Springer Verlag.

11 [14] Turban, Aronson, (2001). Decision Support Systems and Intelligent Systems, 6th Ed., Prentice Hall.