NEURAL NETWORKS

Tom Heskes
SNN, University of Nijmegen, The Netherlands

David Barber
Aston University, Birmingham, UK

Keywords: neural networks, emergent behavior, statistical physics, learning, statistical inference, frequentist approaches, Bayesian methodology

Contents

Summary

This article gives a biased overview of neural-network research. Neural networks are computation models, inspired by information processing in the brain. Two aspects of neural networks are emphasized: the ability to generate complex behavior from many simple processing elements and the property to learn from data.

Emerging complex behavior is most prominent in classical neural networks like the Hopfield network and the Boltzmann machine. Techniques from statistical physics have been applied to understand their operation.

The learning capability of neural networks forms the basis of most practical applications of neural networks. In this context, neural networks have shifted from heuristic black boxes to advanced statistical tools. Learning can be interpreted as a statistical inference problem, which can be solved either using frequentist or Bayesian approaches.

Neural networks have been looked at from many different viewpoints. The cross-fertilization between neural networks and many other fields will continue to grow and improve neural networks, both as advanced statistical tools for solving practical problems and as computational models for understanding how the brain works.

1. What are neural networks?

The field of neural networks covers a large area ranging from theoretical neurobiology to statistical physics and machine learning. Covering in depth such a wide range of topics would be beyond the scope of this article, in which we prefer to give instead some insights into issues related to the history of the field, theoretical concepts that underpin the field, and then give an overview of more practical applications. There are many good textbooks available, covering much of the material in this review, see the references for details.

What is it that ties such seemingly disparate areas of research together? In all these areas there is a common interest in the properties of a (possibly very large) number of relatively simple processing units (neurons/spin-particles/elementary computing units) when they are coupled together. It is the belief that the emergent phenomenon when such simple units are coupled together is capable of explaining such complex effects as intelligence, memory, magnetism and can provide a useful basis for machine learning and computation.

In such an interdisciplinary field, neural networks can take on somewhat different meanings, and the demands or emphasis of the research are correspondingly different. Following Churchland and Sejnowski (and adding the “Physicist” viewpoint), we can loosely categorize the demand/complaint pairs of different researchers, based on this underlying neural-network research as

The Neuroscientist

Show me results of neuromodeling that help explain or predict experimental results. Non neuroscientists do not know anything much about neuroscience even though they are doing “neural modeling”.

The Psychologist

Show me results of neuromodeling that help explain or predict psychological functions and behaviour.

Non psychologists do not know anything much about the results from psychophysics and psychology even though they are modeling psychological capacities and performance.

The Computer Scientist

Show me results of neuromodeling that help understand the nature of computation and representation or that yield new ideas about these things.

Non computer scientists do not know anything much about electrical circuits, mathematical analyses, or existing theories of computation.

The Philosopher

Show me results of neuromodeling that are relevant to philosophical problems concerning the nature of knowledge, the self and the mind.

Non philosophers do not understand some of the useful, time saving, and agony-saving contributions of philosophers in constraining questions about how the mind works.

The Physicist

Show me results and insights from neuromodeling that demonstrate how the macroscopic behaviour of complex systems can be understood to be dependent on a compact description of the system. Non physicists do not know anything much about statistical mechanics and how to bridge the gap between a microscopic description and understanding the resulting macroscopic behaviour.

We do not proceed here to attempt to unify these differing viewpoints. Indeed, this diversity of viewpoints is, in our opinion, a healthy feature. Our own work has been most closely associated with the "computer scientist" and "physicist" viewpoints. As we see it, the problem of making artificial machines endowed with some of the functionality of biological organisms (*e.g.,*visual processing of information) is so complex, and the *a priori* space in which to search for possible solutions so vast, that we must turn to those biological organisms and their specific functionality if we are to succeed in our goal. In so doing, we need to familiarise ourselves with relevant mathematical frameworks since, ultimately, we wish to find a mathematical description of the procedure. In spirit this is similar to David Marr's view of artificial intelligence, which relates artificial intelligence to coding specific biological functions. (See his article in Boden's Book and others therein for an introduction to different viewpoints on artificial intelligence and related philosophical issues). In concentrating on making an "artificial neural network" which is capable of information processing in a manner similar to the brain, there are some observations about the brain which highlight some of the differences to conventional computation (see Hertz, *etal*):

- It is robust and fault tolerant. Nerve cells die every day without affecting its performance significantly
- It is flexible. It can easily adjust to a new environment by "learning" – it does not have to be programmed in a standard computer language
- It can deal with information that is fuzzy, probabilistic, noisy, or inconsistent.
- It is highly parallel
- It is small, compact, and dissipates very little power

Ultimately, we would like to be able to make a machine that can perform certain information processing tasks such as face or speech recognition that humans can do with consummate ease. The starting point here is that conventional approaches based on Artificial Intelligence have reached an impasse, since the task of formally specifying a task such as face recognition is either unclear or too complex to be handled in a conventional way.

We split our review into the following main sections. The first, section (2) deals with the neural biological background that ultimately motivated the field. We then show how this motivated some of the earliest artificial neural network models in section (3).) The generalisation of such early models leads us into the realm of statistical physics in section (4). This subfield of neural networks focuses on the emergence of macroscopic behaviour from the detailed microscopic desciptions of neural networks. A particularly useful area of research has been the development of neural networks as advanced statistical models. In section (5) we review some of the progress that has been made in the machine learning community under the general banner of neural networks and in particular those developments in learning non-linear mappings parameterised as perceptrons. This section contains more immediate practical issues of neural networks and related methods. There is a wealth of material on this topic, and an introduction can be found in the references. More general and recent discussions on how to train models such as neural networks are treated in section (6), where we compare frequentist and Bayesian approaches. How neural networks have found commercial success in applications is outlined in section (7). Finally, we conclude in section (8) with a summary and outlook on where artificial neural networks might be heading in the near future.

2. Neurobiology

2.1 Neurons

We follow our references in this overview of neurobiology. Neurons are the basic structural components of the brain. A neuron is an individual cell, specialised by architectural features that enable fast voltage changes across its membrane as well as voltage changes across neighbouring membranes. Brains are assemblies of such cells, and while an individual neuron does not see or reason or remember, brains do. How can we get from ion movement across cell membranes to memory or perception in brains. What is the nature of neuron-neuron connectivity and interactivity? What makes a clump of neurons a nervous system? Two ground-breaking discoveries in the nineteenth century established the foundations for a science of nervous systems. (1) Macro effects displayed by the nervous systems depend on individual cells, whose paradigm anatomical structure include both long tails (axons) for sending signals and treelike proliferation (dendrites) for receiving signals, see Figure 1. (2) These cells are essentially electrical devices; their basic business is to receive and transmit signals by causing and responding to electric current. Within the last few decades, an enormous amount has been learned about neurons: about their electrophysiology, microanatomy, connectivity and development. If we know so much about the fundamental microscopic aspects of the brain, neurons, surely we also have a good understanding of macroscopic aspects such as how the visual or motor system works. In fact, we do not. It could be that we simply do not yet understand in enough detail how neurons work - ultimately, proponents of this bottom-up research approach contest that we will be able to understand large scale phenomena in the brain.

However, the main argument of some theoretical neurobiologists is that, no matter what level of detail the individual neuronal aspects of the brain are understood, this is not sufficient to explain the complex large scale properties of the brain, such as visual awareness. Such researchers contend that such complex behaviours can only be realised when such neurons and coupled together, producing a dynamic, highly non-linear information processing system, the power and properties of which cannot be understood merely by the study of neurons in isolation. This approach is central to the field of computational neuroscience. This field aims for biological realism in computational models of neural networks which may, however, study at times relatively simple models to see if they are sufficient at qualitatively explaining emergent biological phenomena.

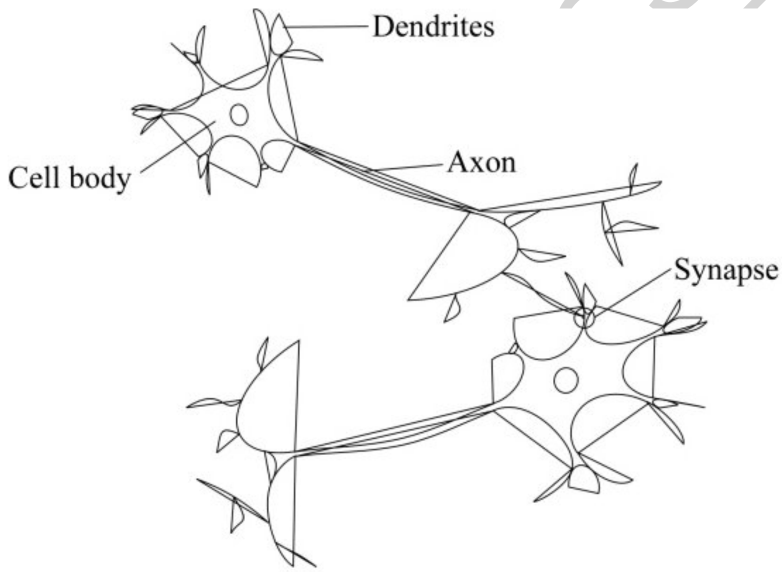

Figure 1: schematic representation of two neurons and their connection point at the synaptic junction. The cell body receives (electrical) input along its dendrites. Provided that this combined input is high enough, a spike or pulse is transmitted along the axon, branching out to many (typically of the order of a thousand) synaptic junctions. These signals are then received by an afferent neuron along its dendrite.

2.2 Simple Neuron Models

The brain is composed of about 10^{11} neurons of many different types, a common class of which has the form depicted in Figure 1. Tree like networks of nerve fibre called dendrites are connected to the cell body or soma, where the cell nucleus is located. Extending from the cell body is a single long fibre called the axon, which branches into strands and substrands. At the ends of these are the transmitting ends of the synaptic junctions, or synapses to other neurons. The receiving ends of these junctions on other cells can be found both on the dendrites and on the cell bodies themselves. The axon of a typical neuron makes a few thousand synapses with other neurons.

The transmission of a signal from one cell to another at a synapse is a complex chemical process in which specific transmitter substances are released from the sending side of the junction. The effect is to raise or lower the electrical potential inside the body of the receiving cell. If the cell body potential of a neuron (after receiving inputs from its neighbours) reaches a threshold, a pulse or action potential of fixed strength and duration is fired, which is propagated along the axonal arborization to synaptic junctions of other cells. After firing, the cell has to wait for a time called the refractory period before it can fire again.

McCulloch and Pitts in 1943 proposed a simple model of a neuron as a binary threshold unit. Specifically, the model neuron computes a weighted sum of its inputs from other units, and outputs a one or a zero according to whether this sum is above or below a certain threshold:

$$\mathrm{n_i}(\mathrm{t}+1) = \Theta\left(\sum_{\mathrm{j}} \mathrm{w_{ij}} \mathrm{n_j}(\mathrm{t}) - \mu_{\mathrm{i}}\right).$$

Here n_i is either 1 or 0, and represents the state of neuron i firing or no firing respectively at time t. Θ is the step function, $\Theta(x) = 1$ if $x \geq 0$ and $\Theta(x) = 0$ otherwise.

The weight w_{ij} represents the strength of the synapse connecting neuron j to neuron i. It can be positive or negative corresponding to an excitatory or inhibitory synapse respectively.

Though individually simple, a collection of McCulloch-Pitts neurons forms a computationally powerful device. Indeed, a synchronous assembly of (sufficiently many) such neurons is capable of universal computation, programmable by choosing weights w_{ij}, and can thus perform any computation that an ordinary digital computer can do.

This simple description differs from real neurons in some fundamental ways.

- Real neurons respond to their input in a continuous way. However, the non-linear relationship between the input and the output is a universal feature. A working hypothesis is that nonlinearity is essential, though not its specific form.
- Real neurons perform a nonlinear summation of their inputs.
- A real neuron produces a sequence of pulses, not a simple output level. Representing the firing rate by a single number n_i, even if continuous, ignores the possibility that pulse phase, the timing of individual “spikes”, not just the rate, encodes a significant amount of relevant information.
- The amount of transmitter substance released at a synapse may vary unpredictably.

A simple generalisation of the McCulloch-Pitts neuron which includes some of these features is

$$n_i(t+1) = g\left(\sum_j w_{ij} n_j(t) - \mu_i\right)$$

where *g* is a continuous function. To take into account some of the stochastic effects, we could alternatively consider

$$p(n_i(t+1) = 1) = g\left(\sum_j w_{ij} n_j(t) - \mu_i\right) \qquad (1)$$

where *g* is a function between 0 and 1, so that (1) represents the probability that neuron *i* fires in a unit time interval.

In most applications of (classical) neural networks, the former interpretation of neurons is applied. That is, the output of each network is a deterministic (nonlinear) function of its inputs. This is then fed successively into other neurons, and the process repeated. We shall mainly deal with this approach in section (5). The stochastic case, in which we consider the output as representing the probability that the neuron fires, is more closely related to systems in statistical physics, and we shall deal with this more closely in section (4). The tools of statistical mechanics may be applied to analysing the properties of both kinds of deterministic and non-deterministic systems. Interestingly, the stochastic model of neurons is a special case of a wider class of statistical models known as graphical models. Graphical models were introduced in response to the failure of traditional expert systems to cope with uncertainty, and is currently a hot research area.

-
-
-

Bibliography

Anderson, J. and E. Rosenfeld (Eds.) (1988). *Neurocomputing: Foundations of Research*. Cambridge: MIT press. [Reprint of the classical papers in the field with new introductions]

Bishop, C. M. (1995). *Neural Networks for Pattern Recognition*. Oxford University Press. [More mathematical oriented introduction which focusses on one application area]

Boden, M. (1990). *The Philosophy of Artificial Intelligence*. Oxford University Press.

Castillo, E., J. M. Gutierrez, and A. S. Hadi (1997). *Expert Systems and Probabilistic Network Models*. Springer. [Discusses methods for incorporating uncertainty into models. This book is gives up-to-date survey.]

Churchland, P. S. and T. J. Sejnowski (1992). *The Computational Brain*. MIT press. [*The Computational Brain* is the first unified and broadly accessible book to bring together computational concepts and behavioral data within a neurobiological framework. The authors address the foundational ideas of the emerging field of computational neuroscience, examine a diverse range of neural network models, and consider future directions of the field]

Gilks, W. R., S. Richardson, and D. J. Spiegelhalter (1995). *Markov Chain Monte Carlo in Practice*. Chapman and Hall. [Emphasis is placed on practice rather than theory, although fundamental theoretical concepts are discussed.]

Hertz, J. Krogh, A. and G. Palmer (1991). *Introduction to the theory of Neural Computation.* Redwood City California: Addsion-Wesley. [Popular textbook for students with a background in physics]

Loredo, T. (1990). From Laplace to Supernova SN 1987A: Bayesian Inference in Astrophysics. In P. Fougere (Ed.), *Maximum Entropy and Bayesian Methods*, pp. 81-142. Kluwer.

MacKay, D. (2000). *Information Theory, Inference and Learning Algorithms*. Cambridge University Press. [A classic, covering Shannon's fundamental theorems to the postmodern theory of LDPC codes].

Biographical Sketches

David Barber completed the University of Cambridge mathematics tripos in 1990. Following a year at Heidelberg University, he studied for a Master's degree in Neural Networks at King's College, London. His PhD on the statistical mechanics of machine learning was carried out in the theoretical physics department at the University of Edinburgh in 1995, followed by spells in Nijmegen, Holland and Aston University. He currently works in the Division of Informatics, Edinburgh University, where his main interests include machine learning, Bayesian methods, graphical models, and links with statistical mechanics.

Tom Heskes received both the M.Sc. and the Ph.D. degrees in physics from the University of Nijmegen, The Netherlands, in 1989 and 1993, respectively. After a year postdoctoral work at the Beckman Institute in Champaign-Urbana, Illinois, he (re)joined the Dutch Foundation for Neural Networks (SNN) in 1994. Currently he is the director of the company SMART Research BV and an assistant professor at the University of Nijmegen. His research interests include theoretical and practical aspects of neural networks and related techniques, with an open eye towards applications.