1990

PANEL 16 BUSINESS APPLICATIONS OF NEURAL NETWORKS: PROBLEMS AND OPPORTUNITIES

Shimon Schocken
New York University, USA

PANEL 16

BUSINESS APPLICATIONS OF NEURAL NETWORKS: PROBLEMS AND OPPORTUNITIES

Panel Chair: **Shimon Schocken**, New York University, USA

PANELISTS: **John Hertz**, Nordita and Neurotech, Denmark
Ludwik J. Liszka, Swedish Institute of Space Physics, Sweden
Douglas L. Reilly, Nestor Inc., USA
Leo Vercauteren, State University of Ghent, Belgium

BACKGROUND

The term "neural network" means different things to different people. Neural networks are popularly perceived as descriptive models, designed to mimic the mechanism of the human brain. Most researchers, however, view the similarity of the neural network architecture to the brain circuitry as no more than a useful analogy. This is because neural networks are powerful computational models in their own right, *regardless* of their biological justification. They offer novel solutions to many problems which defy standard algorithmic techniques and they lend themselves nicely to new developments in parallel and optical hardware.

During the past decade, neural networks were constructed to carry out a wide variety of computational tasks, primarily in scientific and engineering applications. Most of this work, however, is experimental; it remains to be seen whether or not neural networks will become a widely-used computational paradigm. Notwithstanding its slow assimilation in the field, the neural networks area is presently bursting with research activity, attracting scores of people from computer science, neuroscience, physics, and psychology. Since the conception of the idea in the mid-1940s, many architectures and training algorithms were developed. A partial list of several major paradigms and their respective developers includes *perceptions*, *feed-forward networks*, *recurrent networks*, *adaptive resonance theory*, and *feature maps*.

Several hardware and software vendors presently offer off-the-shelf products which implement some of these models. The typical product is a software shell that enables end-users to develop and test neural networks in the context of specific problems. Is it possible that these programs will eventually become popular tools, like statistical packages and spreadsheet programs? The answer seems to go far beyond user-interface and performance considerations; it is related to deeper issues regarding the appropriateness of the neural network approach to the special nature of business applications. Hence, we take the position that neural networks offer a great deal of promise, as well as a certain degree of uncertainty. In this panel, we wish to explore the limitations and potential of this technology in the context of business applications.

WHAT *IS* A NEURAL NETWORK?

Broadly speaking, a neural network consists of three things: a task, an architecture, and a training algorithm. The task specifies what the network is supposed to do: convert hand-written text into an ASCII file, classify sonar images onto submarine profiles, or determine the credit-worthiness of personal loan applications. These problems seem to be very different from each other. Surprisingly, it has been shown that one can successfully model all of them using generic neural architectures and domain-free training algorithms.

The neural architecture consists of a set of processors which are linked to each other, forming a network topology. Each processor is capable of computing a simple function and transmitting the result of its computation to many other processors. The strength of these messages is modulated by a set of weights which parameterize the inter-processor connections. The topology of the network, and the functions that the processor compute, are specified by the network designer a priori; they become a fixed part of the network's architecture. In contrast, the weights of the inter-processor connections are data-driven; they are continuously adjusted by a mathematical procedure which is typically referred to as a *training algorithm*. Many training algorithms exist and all of them are related in one way or another to statistical error-minimizing procedures.

To illustrate, consider a neural network designed to fine-tune an expensive direct-mail campaign. Our objective is to classify the company's customers into three classes: *high*, *medium*, and *low* priority targets. This classification can be modeled through a neural network architecture consisting of three layers of processors. The input layer, consisting of some fifty processors, is used to represent demographic and financial data items, which are pulled out from the company's file. The output layer consists of three processors which are labeled "high," "medium," and "low." The input and output layers are connected to each other through an intermediate set of processors, designed to capture complex and possibly non-linear interactions among the input and output variables. Each connection is parameterized by a numeric weight. At the beginning of the training session, all of the weights are initially set to small random values.

The training session proceeds by showing the network a set of hundreds or thousands of customers whose true priorities are known ex-post from historical data. The customer's data are fed into the network incrementally, one data-set at a time. The network computes the customer's priority by turning on one of the output processors. Next, a feedback mechanism tells the network whether or not this was a correct classification. In the case of a disagreement, a training algorithm adjusts the weights in an attempt to correct the error. This is a delicate operation, because the new set of weights must be chosen so that previously correct classifications will not be reversed. The training session continues until the network reaches a satisfactory level of performance or until the designer runs out of examples. At that point, the network's performance is tested on the entire training data-set as well as on real data.

PANEL OVERVIEW

The goal of the panel is to explore, demonstrate, and discuss the potential and limitations of neural networks in business applications. In spite of the complexity of the underlining subject, the panelists will attempt to minimize the use of technical clutter and stay away from esoteric terminology. This is based on the assumption that every equation or technical term is likely to decrease the audience's interest in the subject matter.

The panel will commence with a mini-tutorial covering the history, technology, and state-of-the-art of the neural networks field. This introduction will set the stage for the second part of the panel, in which we will present neural network solutions to problems in finance, marketing, and operations management. Specifically, we will survey the use of neural networks in such applications as credit scoring, market analysis, prediction, classification, queuing, and scheduling.

In the final part of the panel, the panelists will draw on their collective work experiences to propose specific recommendations to companies who are presently considering the use of neural networks technology. Each panelist will discuss such issues as face-validity, accountability, robustness, user interface, and management acceptance.

The panel includes researchers, practitioners, and vendors of neural networks and represents five different countries. The panelists have extensive experience in developing and testing neural networks in realistic applications. Together, they offer a very broad perspective on the state of the art of the field in Western Europe and in the United States.