CONF-890545--8
DO Note 848

DOE/ER/3130-47C

THE USE OF NEURAL NETWORKS IN THE D0 DATA ACQUISITION SYSTEM† *

David Cutts, Jan S. Hoftun, Andrew Sornborger
Brown University
Providence, RI 02912

DOE/ER/03130--47C

DE89 013960

Richard V. Astur
Michigan State University
East Lansing, MI 48824

Christopher R. Johnson, Raymond T. Zeller
ZRL
Cranston, RI 02905

Abstract

We discuss the possible application of algorithms derived from neural networks to the D0 experiment. The D0 data acquisition system is based on a large farm of MicroVAXes, each independently performing real-time event filtering. A new generation of multiport memories in each MicroVAX node will enable special function processors to have direct access to event data. We describe an exploratory study of back propogation neural networks, such as might be configured in the nodes, for more efficient event filtering.

Introduction

Parallelism plays a central role in advanced high energy physics data acquisition systems, from digitization and collection through real-time event filtering in online computer farms. Both the discrete sampling nature of detectors and the event structure of particle physics data lead to an explicit parallelism. As data handling requirements for real-time systems become more severe, we need to exploit to an ever higher degree this natural feature of our data. Although digitization electronics, data paths, and microprocessors are easily replicated for more parallelism, the standard filter algorithms which run in the farm nodes are not so easily handled. Therein lies a major advantage of algorithms based on neural networks: they are intrinsically parallel.

Back Propagation Networks

Neural networks [1] are receiving increasing attention in many fields as a means of deriving or implementing pattern recognition tools [2]. As interest has grown so has the variety of network configurations and characteristics.

Among the more promising of these are "back propagation" networks, which have been shown capable of pattern recognition in diverse applications, from sonar signal analysis to stock market prediction. Our inquiry into the application of neural networks to high energy physics, and specifically D0, has focused on back propagation networks because of their demonstrated recognition ability, the ease with which they can be simulated, and a natural means through which they can be implemented in the D0 data acquisition system.

Neural networks are parallel structures consisting of units (each modelling a neuron) with many interconnections between units, and parameters called "weights" specifying the strengths of each interconnection. As the network learns, it modifies these weights; and its training encodes features of the pattern in the values of all the weights. Back propagation networks, our specific example, have several layers of units: an input layer, one or more "hidden" layers, and an output layer. Every unit in a given layer is connected to each unit in the succeeding layer, and the flow as the network responds to an input is in this "feed-forward" direction. All the information used by the network, for each event, is presented in the form of a vector, with each element linearly scaled by an input unit, to form an output vector. The output vector of a given layer is multiplied by the matrix of the weights between this layer and the next, to produce an input vector to the next layer. Each unit of the 2nd and subsequent layer uses a "sigmoid" activation function [1] to translate the weighted sum of its inputs into an output. With this combined action of forming weighted sums and evaluating response, the network converts an input vector into an output: in the example to be described, a two-unit vector with (1,0) and (0,1) representing the extremes of recognition. The label "back propagation" describing these networks refers to the technique used in the learning phase whereby the network's response (output vector) to a given input is compared with the desired response, and the difference is propagated back through the network to adjust the weights.

The operation of back propagation networks, described

above (and more explicitly in Ref. 1), lends itself to straightforward simulation and implementation. Our study of the performance of the network relies on a computer model, which can be readily reconfigured, trained with data, and tested. We have used a commercial package "Professional-II" [3] running on an IBM PS/2 (Model 70) for several studies described previously [4]. For the exercise described in this paper, we wrote a simulation package which runs under VAX/VMS, since convenient access to D0 data and computer resources was important.

The D0 Data Acquisition System

Data acquisition at D0 [5,6] is accomplished by dumping the raw data for an entire event (250 KBytes) from the 100 VME digitization crates directly into memories of a MicroVAX system, selected from an array of Level-2 trigger nodes. Data flows to the node from VME over eight data paths having an aggregate bandwidth of 320 MBytes/second. Each node will be equipped with eight channels of ZRL Q22MPM multiport memories, which receive the event; as these memories are directly accessible to the processor as well as to a high speed output channel, additional moves of data should not be necessary for the Level-2 filter package to operate. As described in reference 5, a special function port on these boards will allow access to the event data by specialized processors. One such implementation, of a special purpose array processor, is shown in Figure 1. This device would lend itself naturally to performing algorithms derived from the back propagation neural networks discussed above, consisting as they do of a few vector-matrix multiplications and vector lookup operations.

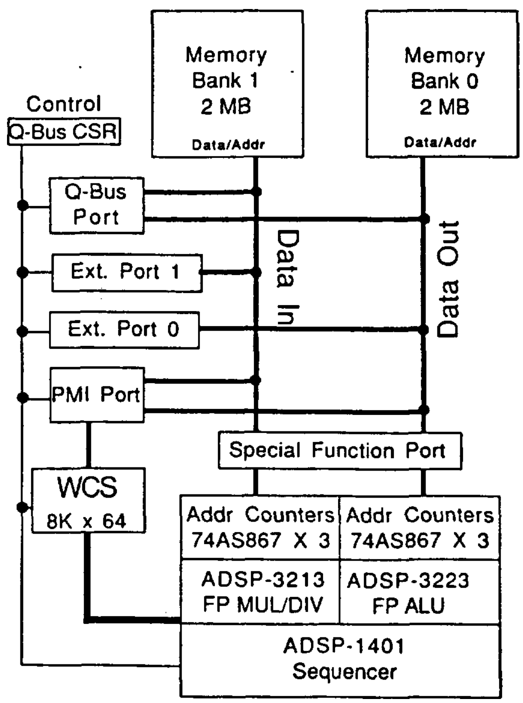


Fig. 1 Multiport memory with special purpose array processor for DØ

An Exercise: Electron Identification at D0

To explore the possibilities of using neural network-derived algorithms at D0, we have studied a particular application, electron identification from calorimeter data. A large sample of Monte Carlo data, simulating the detection of proton-antiproton collision events with the D0 detector, was available. This data, generated using the GEANT simulation package, was produced at Brown two years ago, using a farm of dedicated MicroVAXes [7]. We extracted from candidate electron showers in this data, specific information from the simulated energy deposits in the D0 uranium-liquid argon calorimeter. For each candidate shower in our sample, we collected the observed deposits in the first five layers in depth and four radial bins about the shower peak, from 0.1 in $(\delta\eta, \delta\phi)$. Typical data for this (4,5) array is shown in Figure 2. We studied both $Z \rightarrow e^+e^-$ data (Fig 2a) and high p_T two-jet background data (Fig. 2b). With each event we included also a tag from the ISAJET event generator which flagged the actual presence or absence of an electron.

a)

b)

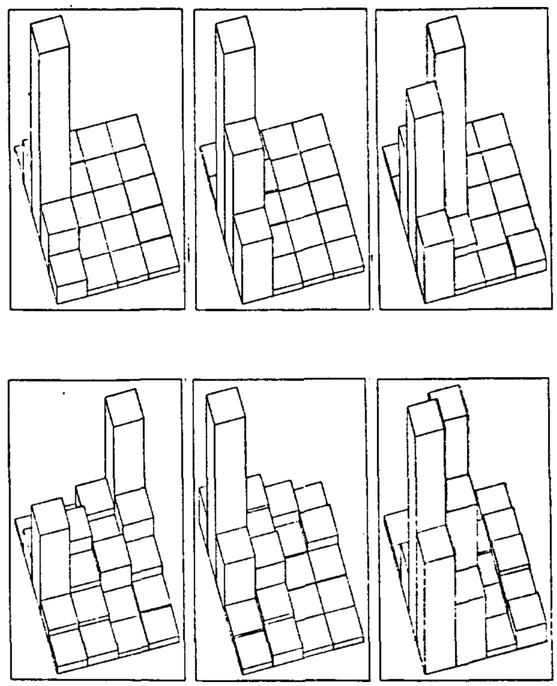
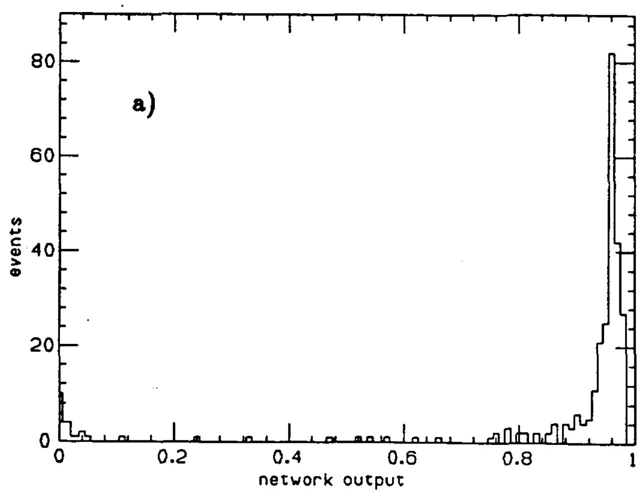


Fig. 2 Sample data of energy deposits in the first five layers and for bins radially about a shower for a) $Z \rightarrow e^+e^-$ events and b) 2-jet background events with $80 < P_T < 120 GeV$. The front central bin is in the lower left corner.

Our model neural network consisted of three layers: an input layer of 20 units (corresponding to the 20 element input vector), a middle layer of 8 units, and an output layer of two units. During training we presented as the desired output the vector (1,0) if the ISAJET eletron tag was present for that event, and (0,1) otherwise. Our simulation package allowed us to vary several parameters related to the learning, including the strength of the error-correction (related to the speed of learning) and the contribution of the correction from the previous event (a "momentum" term) [8]. We found a low but non-zero momentum term, and a gradually decreasing learning strength to be optimum.



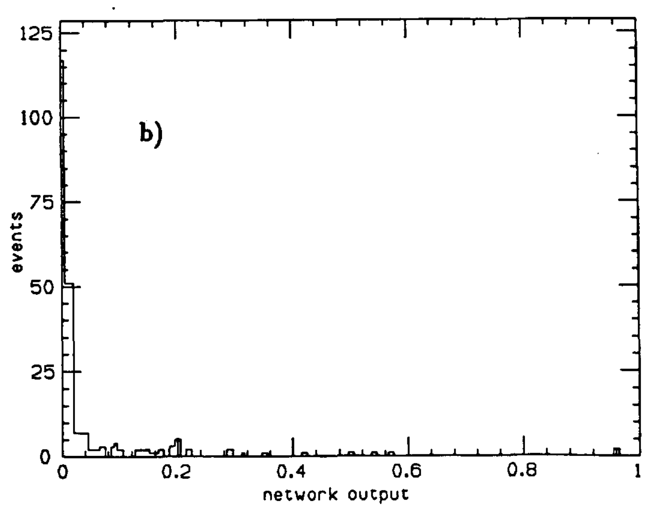


Fig. 3 Response of the network to showers from a) $Z \rightarrow e^+e^-$ events and b) 2-jet background events ($80 < P_T < 120$).

Having trained our network simulation on one set of data, a mixture of both $Z \rightarrow e^+e^-$ and 2-jet data, we tested it on a similiar but independent data set. The results of this test are shown in Figure 3, where the response of the network to events tagged as electrons by ISAJET (Fig. 3a) is clearly distinct from its response to events without electron tags (Fig. 3b). We have also tested the network with backgound data generated in lower intervals in p_T, which are confused more often with electrons, and the results are included in Table 1. Shown here are the numbers of events considered "electrons", with a selection on the network output that passes 90 percent of the sample tagged as electrons. For comparison, a selection based on a number of cuts on the ratios of transverse and radial energies in the event produces a poorer rejection of background data, while at the same time only accepting 75 percent of the true electron sample [9].

Table 1.
Comparison of neural network and standard of algorithms

	Neural Network		Standard Algorithm	
data set	# events in sample	% recognized as "electrons"	# events in sample	% recognized as "electrons"
$Z \rightarrow e^+e^-$	275	90.2	643	75
2 jet ($80 < P_T < 120$)	225	0.9	826	5
2 jet ($40 < P_T < 80$)	370	1.6	681	10
2 jet ($20 < P_T < 40$)	122	8.2*	104	23

* This becomes 0.8%, for 80% recognition of $Z \rightarrow e^+e^-$ electrons.

Summary

Neural networks used to develop pattern recognition algorithms force a parallel solution. This explicit parallelism of such algorithms is the key to their high speed implementation for high energy physics pattern recognition. Such algorithms, embedded in a natural way in the D0 data acquisition system, are a promising addition to D0's event filtering capabilities.

References

[1] For a basic reference to neural networks, see David Rumelhart et al., Parallel Distributed Processing, MIT Press (1986).

[2] An early application to high energy physics is described in B. Denby, "Neural Networks and Cellular Automata in Experimental High Energy Physics", Comp. Phys. Commun. 49, p. 429 – 448 (1988).

[3] NeuralWare Inc., 103 Buckskin Court, Sewickley, PA 15143, USA.

[4] D. Cutts et al., "Neural Networks for Event Filtering at D0", International Conference on Computing in High Energy Physics, Oxford (April 1989), to be published.

[5] D. Cutts et al., IEEE Transactions on Nuclear Science, 36, p. 738.

[6] D. Cutts et al., "Data Aquisition at DØ", International Conference on Computing in High Energy Physics, Oxford (April 1989), to be published.

[7] D. Cutts, J. S. Hoftun, "Running a Large Monte Carlo Program in a Farm of MicroVAX-II Computers under VAXELN", presented at Argonne Workshop on Detector Simulation for the SSC, (August 1987).

[8] See Ref.1, Chapter 8 for a discussion of the learning strength and momentum coefficients.

[9] The comparison algorithm is representative of the performance of such methods but is not optomized. We are also investigating another approach, the "H-Matrix", which has been used on other D0 data and is known to give superior results (see R. Englemann et al., NIM 216, 45 (1983)).

† This work is supported in part by the U.S. Department of Energy.

* Presented at the conference REAL-TIME'89 Williamsburg (May 1989), to be published.