Neural network real time event selection for the DIRAC experiment

P. Kokkas, M. Steinacher, L. Tauscher, S. Vlachos*

Institute for Physics, University of Basel, Klingelbergstr. 82, CH-4056 Basel, Switzerland

Abstract. The neural network first level trigger for the DIRAC experiment at CERN is presented. Both the neural network algorithm used and its actual hardware implementation are described. The system uses the fast plastic scintillator information of the DIRAC spectrometer. In 210 ns it selects events with two particles having low relative momentum. Such events are selected with an efficiency of more than 0.94. The corresponding rate reduction for background events is a factor of 2.5.

INTRODUCTION

The DIRAC experiment (1) at CERN, is designed to measure the life-time of an atom consisting of two oppositely charged pions, henceforth called pionium. This is achieved by detecting the two charged pions emerging from the pionium's breakup.

Pioniums are produced when the CERN-PS 24 GeV/c proton beam hits a thin solid target. When they interact electromagnetically with matter, they break up into two oppositely charged pions with very small relative momentum ($Q < 3$ MeV/c) and, in particular, a very small angle between them (typically less than 0.3 mrad). The DIRAC detector is designed to measure the momenta of the two pions with highest accuracy. This is achieved with a double arm magnetic spectrometer.

The neural network model developed for the DIRAC first level trigger application uses a pair of hits downstream the DIRAC magnet (hits in the Ionization Hodoscope, IH) together with a hit in each spectrometer arm (Vertical Hodoscopes, VH hits). There are two separate Ionization Hodoscope detection planes each with 16 vertical slabs. Hence each IH plane is used independently in conjunction with the same downstream detectors, and the two event decisions are eventually combined together with a logical OR function. This trigger system is called DNA (Dirac Neural Atomic trigger).

* e-mail: S.Vlachos@cern.ch

DATA SAMPLES AND FORMAT USED

In what follows events that have to be selected by DNA are called 'good', and those to be rejected 'bad'. The definition of good events is the following:

- $Q_x < 3$ MeV/c (Q along the x, horizontal direction),
- $Q_y < 10$ MeV/c (Q along the y direction, direction of the magnetic field),
- $Q_l < 30$ MeV/c (longitudinal Q component).

Bad events are the ones not satisfying any of the above criteria. In total we have used 10000 events of each kind for the neural network training phase. An equal number of events has been used to estimate the DNA performance during development, before it was actually implemented.

The DNA trigger processes events with a maximum of 2 hits in each VH and a maximum of 5 hits in each IH layer. Events with more hits are rare and not evaluated by DNA (as combinatorics resolution is not feasible online). However for improved signal acceptance all such events are read-out for further off-line analysis.

Since there may be more than 2 hits in a IH layer, DNA uses only the pair of hits that are closer together. These are the ones that most probably correspond to the two charged pions of a decayed pionium. In case of a single hit in a IH layer, it is considered that both charged particles went through the same IH slab, and hence this slab is used twice.

The VHs can have one or two hits each. DNA forms all four possible combinations with a single hit in each VH. Hence four different input patterns are generated per IH layer per event. Each such pattern contains the same IH information. Four independent but identical neural networks evaluate further these input patterns (one per VH combination).

THE NEURAL NETWORK

For DNA we have used a feed-forward neural network model based on error back propagation (for a general description of neural network algorithms see for example (2) and (3)). It is well established that a feed-forward neural network with one hidden layer can represent any function of the input data. Consequently the DNA neural network has 18 input nodes, one hidden layer with two nodes and one node in the output layer. This model has been initially developed for a completely different application (4).

A simple but crucial choice we had to make was the selection of the appropriate transfer function for each neural network node. In order to avoid problems related with the special treatment of 0 in training algorithms, we have used instead a 'tanh' activation function (which has a sigmoid form from -1 to 1 for an input range from -1 to 1).

Using the data samples described in the previous section we have run the training algorithm for 2000 epochs (an epoch consists of presenting the whole training data sample once to the neural network, and obtain consequently one modification of all the weights).

The performance of the trained neural network is shown in Fig. 1. As all events with a neural network output above a given threshold are tagged as good and are accepted, the efficiency of the selection as a function of the threshold value can be established. Similarly one can establish the efficiency of selecting background. Fig. 1 shows these two efficiencies one against the other for various values of the neural network's threshold.

All four combinations for each of the two IH layers are evaluated in parallel by the same neural network developed above. Hence this network is used eight times in the DNA system. If any of these neural networks gives an answer (output value) greater than a common predetermined threshold, the event is accepted.

HARDWARE IMPLEMENTATION

The DNA application is based on a custom built hardware implementation of the neural network algorithm, originally developed for the CPLEAR experiment (5). The flexibility of this implementation allowed eventually the system to be incorporated in the L3 first-level trigger system (6) and now in DIRAC. A comparative study (7) has shown that more complicated algorithms or hardware architectures are in general less performant. A single card executes a feed-forward neural network with 1 single hidden layer with two nodes in 60 ns. This performance is obtained by using RAM memories with preloaded contents to perform the most computationally demanding part of the algorithm.

The complete DNA trigger consists of two identical chains, one for each IH detector plane. Each DNA chain is based on three separate electronics modules: The Interface and Decision card (IFD) (8, 9), the neural network (NN) cards (10) and a power-PC VME master CPU card (Motorola MVME–2302). The IFD card has been developed for the DNA trigger, the NN-cards are identical to the ones used earlier. Only the CPU-card is commercial (for ease of technical support and software development). The IFD card feeds four independent and identical NN cards (running the same NN algorithm), corresponding to the four combinatorial VH possibilities.

Since April 2000, the DNA trigger is operating in the DIRAC experiment as an integral part of the whole trigger system. Each event is fully classified in less than 210 ns.

EVENT SELECTION AND REAL-TIME PERFORMANCE

In order to study the DNA performance, data with DNA in flagging mode have been analysed. In that mode DNA only flags events as good or bad without rejecting any, so that events normally rejected by DNA can also be studied. In the following the threshold on the neural network output has been set to 0.0 (i.e. in the middle of the neural network output range).

The system's acceptance as a function of Q_{tot} is shown in Fig. 2. The value at very low Q_{tot} gives the DNA acceptance for good events. The overall DNA reduction rate can also be obtained from this figure by convoluting this distribution with the Q_{tot} distribution of two track events in DIRAC.

The performance of the DNA system can be summarized in the following two numbers:

- Acceptance of low Q_{tot} pion pairs: 0.94, derived from Fig. 2,
- Background rate reduction: 2.5, measured during data-taking.

Note here that the above numbers are higher than the ones concerning the neural network performance as per event up to four possible hit combinations are used and if any of them gives a high neural network output (above threshold) the event is accepted.

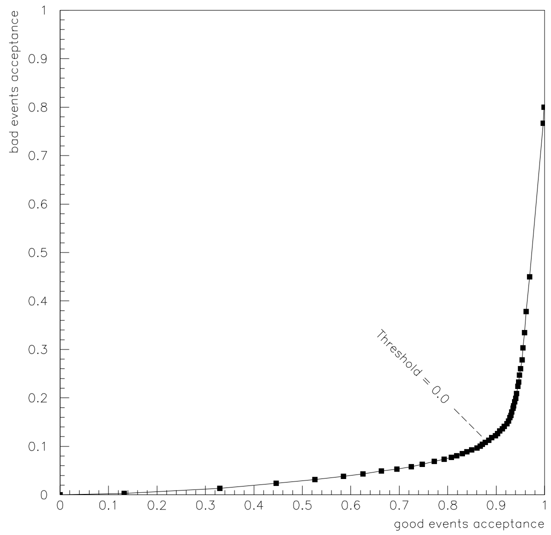

FIGURE 1. The relation between the signal acceptance (good events) and background acceptance (bad events). Each point corresponds to a different neural network threshold. The standard DNA working point (threshold = 0.0) is also indicated on the figure for completeness.

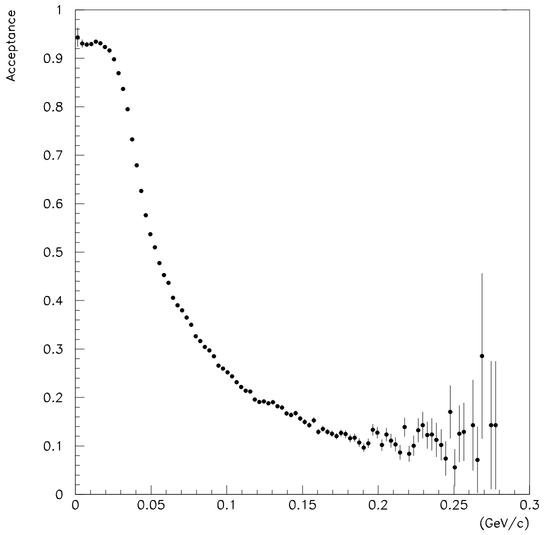

FIGURE 2. The efficiency of DNA as a function of Q_{tot}.

CONCLUSIONS

The DNA trigger is successfully used in the DIRAC experiment since it was first installed. It reduces the event rate by a factor of 2.5 while keeping more than 0.94 of low relative momentum pionic pairs. The hardware implementation based on custom built electronics allows a decision time of only 210 ns. No particular bias is introduced by the DNA event selection.

REFERENCES

1. B. Adeva et. al. Lifetime measurement of $\pi^+\pi^-$ atoms to test low energy QCD predictions, CERN/SPSLC 95–1 (1995).
2. J. Hertz et al., Introduction to the theory of neural computation Addison-Wesley, New York 1991.
3. C. Bishop, Neural networks for pattern recognition, Claredon press, Oxford 1995.
4. G. Athanasiu et. al., Nucl. Instr. and Meth. A324, 320 (1993).
5. F. Leimgruber et. al., Nucl. Instr. and Meth. A365, 198 (1995).
6. D. Haas et. al. Nucl. Instr. Meth. A401, 19 (1999).
7. S. Vlachos et. al., Nucl. Instr. and Meth. A385, 361 (1997).
8. S. Vlachos, 'A global overview of the Basel neural network project after 10 years of experience', 'Sixth International Workshop on Software Engineering, Artificial Intelligence Neural Nets, Genetic Algorithms, Symbolic Algebra, Automatic Calculation (AIHENP99)', 12–16 April 1999, Heraclion (Greece), proceedings to appear.
9. M. Steinacher, 'Interface and Decision Card; The design', Revision 2.0, Internal Report, University of Basel, 9–6–1999.
10. M. Steinacher, 'Hardware implementation of a fast neural network', 'Sixth International Workshop on Software Engineering, Artificial Intelligence Neural Nets, Genetic Algorithms, Symbolic Algebra, Automatic Calculation (AIHENP99)', 12–16 April 1999, Heraclion (Greece), proceedings to appear.