AN ON-LINE NON-LEPTONIC NEURAL TRIGGER APPLIED TO AN EXPERIMENT LOOKING FOR BEAUTY

C. Baldanza, F. Bisi, A. Cotta-Ramusino, I. D'Antone,
L. Malferrari, P. Mazzanti, F. Odorici, R. Odorico, M. Zuffa
INFN ANNETTHE, Bologna, Italy

C. Bruschini, P. Musico, P. Novelli
INFN, Genova, Italy

M. Passaseo
CERN, Geneva, Switzerland

Presented by R. Odorico
E-mail: odorico@bo.infn.it

ABSTRACT

Results from a non-leptonic neural-network trigger hosted by experiment WA92, looking for beauty particle production from 350 GeV π^- on a Cu target, are presented. The neural trigger has been used to send on a special data stream (the Fast Stream) events to be analyzed with high priority. The non-leptonic signature uses microvertex detector data and was devised so as to enrich the fraction of events containing C3 secondary vertices (i.e. vertices having three tracks with sum of electric charges equal to +1 or -1). The neural trigger module consists of a VME crate hosting two ETANN analog neural chips from Intel. The neural trigger operated for two continuous weeks during the WA92 1993 run. For an acceptance of 15% for C3 events, the neural trigger yields a C3 enrichment factor of 6.6-7.1 (depending on the event sample considered), which multiplied by that already provided by the standard non-leptonic trigger leads to a global C3 enrichment factor of ~ 150 . In the event sample selected by the neural trigger for the Fast Stream, 1 every ~ 7 events contains a C3 vertex. The response time of the neural trigger module is 5.8 μ s.

The idea of developing signatures for the dominant non-leptonic decays of heavy flavor particles by exploiting correlations within the associated events was put forward some time ago^{1,2)}. The technique proposed at that time was Fisher's linear discrimination³⁾, which corresponds to a simple Feed-Forward Neural Network with one input and one output layer and a single output neuron. The recent appearance of neural microprocessors, able to implement multi-layered neural networks with response times of a few microseconds, has stimulated efforts for the realization of on-line neural triggers capable of exploiting non-leptonic signatures of heavy flavors in high luminosity particle accelerators. Projects of this type were announced at the IEEE 1992 Nuclear Science Symposium⁴⁻⁶⁾. So far, the only hardware development presented consists of a commercially available VME card, ELTEC SAC-711⁷⁾, hosting the ETANN chip⁸⁾ which has been used off-line with simulated events⁹⁾.

The neural trigger module, whose realization and on-line running results are presented here[#], consists of a VME crate utilizing the neural chips ETANN (analog) from Intel⁸⁾ and MA16 (digital) from Siemens¹¹⁾, Fig. 1. Its ETANN part has operated within the WA92 experiment at CERN¹²⁾ during the 1993 run. WA92 is an experiment looking for the production of beauty particles by a π^- beam at 350 GeV/c impinging on a Cu target.

Task of the Neural Trigger was to select events by exploiting a purely non-leptonic decay signature and to accept them into a special data stream, the Fast Stream, containing events to be analyzed with high priority and which previously collected only events with semi-leptonic decay signatures. Specifically, the Neural Trigger was trained to enrich the fraction of events with C3 secondary vertices, i.e. branching into three tracks with sum of the electric charges equal to +1 or -1. C3 vertices represent a natural signature for charm and beauty non-leptonic decays. They cannot be reconstructed on-line (besides the computing time length, the fine-resolution microstrip Decay Detector is too slow to be used for triggering), and thus they cannot be directly used as a trigger signature. In order to enrich the fraction of C3 events in the selected sample, the Neural Trigger exploits correlations characteristic for C3 events among quantities which can be calculated on-line from the hits measured in the low-resolution silicon microstrip Vertex Detector. Such correlations are "learned" by the neural network, loaded on the Neural Trigger, in a previous off-line stage (the training stage) during which the neural network is presented with two separate event samples certified off-line as containing and not containing, respectively, C3 vertices, and an appropriate training algorithm changes the internal parameters of the neural network so that its response becomes as widely different as possible for the two classes of events.

The training of the neural network was done using WA92 data collected in the 1993 run, with C3 events certified off-line by the Trident event reconstruction

[#] A preliminary presentation has been made in [10].

program. The events used had been accepted by the WA92 Standard Non-Leptonic Trigger (SNLT), so that the task of the Neural Trigger was to provide additional C3 enrichment for the events it accepted into the Fast Stream. Input to the Neural Trigger was provided by the Beauty Contiguity Processor (BCP) of Darbo and Rossi^{13,14)}, which determined tracks and their impact parameters on-line, using hit locations in the silicon microstrip Vertex Detector. The BCP output, arranged in five 64-bit hit-maps each one corresponding to separate impact parameter windows, was preprocessed within the neural crate to yield 16 input variables for the ETANN chips. The neural trigger has been in operation for two weeks, during which its stability was verified.

An introduction to artificial neural networks suited for applications in high-energy physics can be found in¹⁵⁾. Here, we present only the basic concepts. Suppose we wish to discriminate from each other two classes of events, A and B, on which two event variables, s_1 and s_2 , are measured (they can represent event shape variables or whatever else), and suppose that the class distributions in the two event variables are those of Fig. 2. If we try to discriminate class A events from class B events by a conventional cut on the single-variable distribution obtained by projecting the entries over the s_1 axis, we will either lose a large fraction of A events when trying to get a pure A sample or tolerate a large contamination when trying to keep a high acceptance for A events. The same occurs when projecting over the s_2 axis. But if we project over the axis labeled as Fisher axis^{3,16)}, the two resulting class distributions do not overlap and, by imposing a cut in the middle, one gets a clean separation without any loss of acceptance for A events. The essential ingredient one is exploiting is that both in A and B events the two variables s_1 and s_2 are correlated and that the correlation is different within the two event classes.

The projection over the Fisher axis, i.e. the Fisher discriminant variable, can be written as

$$a = \sum_k \omega_k s_k + \theta$$

The Fisher discriminant represents the simplest type of Feed-Forward Neural Network, having 1 layer and 1 output neuron, a representing the activation of the neuron. Its internal parameters, the weights ω_k and the threshold θ , can be straightforwardly determined from the centroids and the correlation (or covariance) matrices of the two event classes. One can easily show that if the two class distributions are gaussian shaped and have the same correlation matrices, no other discriminant can outperform the Fisher discriminant. If the distributions are far from gaussian and, in particular, they exhibit concavities (think e.g. of banana-shaped distributions) the simple Fisher discriminant may turn out to be inadequate. A way out is provided by iterating the Fisher construct. Namely, by introducing more axes defining new coordinates (1st layer), and using them again as input to Fisher discriminants of a subsequent layer, after having them filtered by a non-linear transfer function (otherwise one would collapse back to a single Fisher discriminant).

A discriminant of this type is called a Feed-Forward Neural Network. Similarly to the Fisher discriminant, its internal parameters can be determined from training samples of certified events using appropriate training algorithms (e.g. the Back-Propagation algorithm¹⁷⁾).

For the problem at hand it has been found that a 2-layer net does not improve the results obtained with a 1-layer net. Thus a simple Fisher discriminant has been loaded on the two ETANN chips, which were left to work independently. The results we present are based on samples of events collected in September 1993, while the training sample dates back to mid August, and they were chosen a few days apart to check for stability. The neural net loaded was organized so as to have an high response from ETANN for C3 events. After conversion to a 8-bit digital number, the response from the two ETANN chips comes in the form of an integer number ranging from 0 to 255. The trigger bit was activated when the response from the ETANN chosen for triggering (the other one was used for cross checks) was above a preset threshold, chosen on the basis of a tradeoff between acceptance and enrichment for C3 events and also taking into account other considerations, like the trigger rate. Fig. 3 shows the dependence on the trigger threshold of the acceptance and the enrichment for C3 events given by the Neural Trigger applied to events already selected by the SNLT. The figure refers to one of the two ETANN chips, the other one giving consistent results. Moving up the trigger threshold, the C3 enrichment factor increases up to a value of ~ 7 . To better see the tradeoff between enrichment and acceptance, one can remove the horizontal axis and plot C3 enrichment directly versus the corresponding C3 acceptance, as done in Fig. 4. The small deviations in performance, observable when moving from one event sample to another, do not appear to be due to the Neural Trigger itself, which has been verified to be stable by independent tests, but are likely due to slight modifications in the experimental apparatus, e.g. caused by temperature excursions. The overall stability of the Neural Trigger performance rather shows its robustness toward changes in the external conditions.

From Fig. 4 one can read that for a C3 acceptance of 15% one obtains a C3 enrichment factor of 6.6-7.1, depending on the event sample. That is just the factor by which the C3 enrichment already given by the SNLT is boosted up. The global C3 enrichment factor provided by the SNLT and Neural Trigger combination is found to be ~ 150 , with a statistical error of $\sim 30\%$, due to the limited statistics available for interaction trigger runs (i.e. with the SNLT turned off). The corresponding global C3 acceptance is $\sim 1\%$. Its smallness is largely due to the fact that in the off-line event reconstruction C3 vertices are requested to have two tracks with $|\text{IPI}| > 20 \mu\text{m}$ (using the data from the high-resolution microvertex Decay Detector, not usable for triggering), whereas the SNLT (using only data from the low-resolution Vertex Detector) requires two tracks with $\text{IPI} > 100 \mu\text{m}$. In the sample of events accepted

into the Fast Stream by the SNLT times Neural Trigger combination, 1 every ~7 events is found to have a C3 vertex.

The stability of the Neural Trigger was independently verified by sending during out-of burst intervals a prefixed set of inputs to its Interface Board and monitoring responses. To ensure stability care has been taken to stabilize the power supply voltage of the two ETANN chips within 5 mV and their temperature within 1° C.

In conclusion, a first example of neural network trigger has been demonstrated by two weeks of continuous running in experiment WA92. The non-leptonic trigger for heavy flavors so realized gives a performance comparable to that of conventional semi-leptonic triggers.

References

1. R. Odorico, *Proc. of the 1982 DPF Summer Study on Elementary Particle Physics and Future Facilities*, Snowmass, Colorado, 1982, p. 478; R. Odorico, *Phys. Lett.* **120B** (1983) 219; G. Ballocci and R. Odorico, *Nucl. Phys.* **B229** (1983) 1.
2. A. Cherubini and R. Odorico, *Z. Physik C* **47** (1990) 547, **C53** (1992) 139.
3. R.A. Fisher, *Annals Eugenics* **7** (1936) 179.
4. R. Odorico, Invited Talk at the *1992 Nuclear Science Symposium*, Orlando, Florida, 1992, Conf. Record p. 822; *IEEE Transactions on Nuclear Science* **40** (1993) 705.
5. P. Ribarics, Invited Talk at the *1992 Nuclear Science Symposium*, Orlando, Florida, 1992, Conf. Record p. 825.
6. W. Badgett et al., Invited Talk at the *1992 Nuclear Science Symposium*, Orlando, Florida, 1992, Conf. Record p. 841.
7. *SAC-700/800 User's Manual*, ELTEC Elektronik, GmbH, Mainz, Germany.
8. *80170NX Electrically Trainable Analog Neural Network Data Booklet*, Intel Corp. 2250 Mission College Boulevard, Santa Clara, CA 95052-8125, USA.
9. B. Denby et al., *Nucl. Instr. and Meth.* **A335** (1993) 296.
10. C. Baldanza et al. (ANNETTHE Collaboration), *Proc. 3rd Int. Workshop on Software Engineering, Artificial Intelligence and Expert Systems for High Energy and Nuclear Physics*, 1993, Oberammergau, Germany, p. 391.
11. U. Ramacher et al., *Proc. Int. Conf. on Systolic Arrays*, Killarney, Ireland, Prentice Hall (1989), p. 277.
12. M. Adamovich et al., WA92 Collaboration, CERN/SPSC 90-10 (1990), *Nucl. Phys. (Proc. Suppl.)* **B 27** (1992) 251.
13. G. Darbo and L. Rossi, *Nucl. Instr. and Meth.* **A 289** (1990) 584.
14. C. Bruschini et al., *1992 Nuclear Science Symposium*, Orlando, Florida, 1992, Conf. Record p. 320.
15. P. Mazzanti and R. Odorico, *Z. Physik C* **59** (1993) 273.
16. M. Kendall, A. Stuart and J.K. Ord, *The Advanced Theory of Statistics*, Vol. 3, 4th ed., C. Griffin & Co. Ltd., London (1983).
17. D.E. Rumelhart, G.E. Hinton and R.J. Williams, "Learning Internal Representations by Error Propagation", in D.E. Rumelhart and J.L. McClelland (Eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* (Vol. 1), MIT Press (1986), p. 318.

NEURAL VME CRATE

Fig. 1 - Schematic view of the Neural Trigger VME crate

Fig. 2- Example showing how Fisher linear discrimination works.

Fig. 3 - C3 enrichment and acceptance obtained when varying the trigger threshold on the Neural Trigger output from the ETANN_0 chip.

Fig. 4 - Same as in Fig. 3, with C3 enrichment plotted directly versus the corresponding C3 acceptance.