SELF-ORGANIZED MAPS FOR TAGGING B JETS ASSOCIATED WITH HEAVY NEUTRAL MSSM HIGGS BOSONS

A. Heikkinen*, S. Lehti, Helsinki Institute of Physics, Helsinki, Finland

Abstract

B tagging is an important tool for separating the LHC Higgs events with associated b jets from the Drell-Yan background [1]. We extend standard neural network (NN) approach using multilayer perceptron in b tagging [2] to include self-organizing feature maps. We demonstrate the use of the self-organizing maps (SOM_PAK program package) and the learning vector quantization (LVQ_PAK) [3]. A background discriminating power of these NN tools are compared with standard tagging algorithms.

INTRODUCTION

In MSSM the heavy neutral Higgs boson production in association with two b quarks is the dominant Higgs boson production mechanism at large values of tanβ at LHC. These associated b jets can be used to extract the Higgs events from the Drell-Yan Z/γ^* background, for which the associated jets are mostly light quark and gluon jets.

b-tagging

In standard methodology a jet can be identified as a b jet using lifetime based tagging algorithm, which relies on displaced secondary vertices and track impact parameter (IP). Impact parameter is the closest approach of the track trajectory to the primary vertex. Figure 1 demonstrates the case in CMS detector at Large Hadron Collider.

For a review of the main algorithms for inclusive b-tagging based on track IP and secondary vertex, see Refs. [1] and [4].

SELF-ORGANIZING MAPS (SOMS)

In comparison to quite common multi layer perceptrons (MLPs) technique (See for example Refs. [5, 6, 7, 8].), only few authors have reported on SOM (also known as Kohonen network) use in separation of background from signal, though some promising results have been achieved [9, 10, 11, 12]. Yet SOM is can provide computationally more simple algorithm, with learning rate faster than what MLPs have.

The most popular unsupervised neural network algorithm SOM, defines a mapping from n-dimensional input data space onto a regular two-dimensional array of neurons [13, 14]:

* aatos.heikkinen@cern.ch

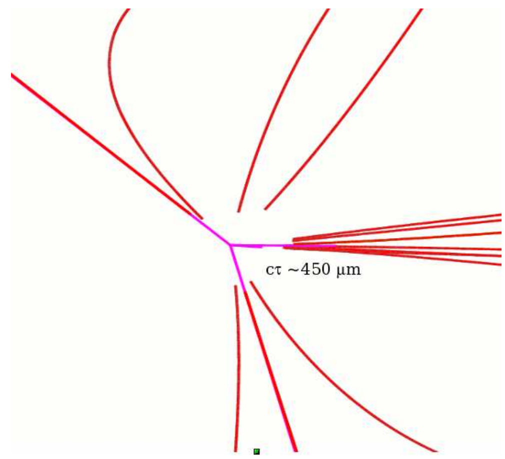

Figure 1: A displaced secondary vertex in a $\mathrm{b\bar{b}H}$ event with $\mathrm{H} \rightarrow \tau\tau$ in CMS. The second b jet is not reconstructed due low jet energy and low track multiplicity.

- Every neuron of the map is associated with an n-dimensional reference vector.
- The neurons of the map are connected to adjacent neurons by a neighborhood relation, which dictates the topology of the map.
- During the unsupervised training phase, the SOM forms an elastic net that folds onto the "cloud" formed by input data and approximates the density of the data.

Competitive process

The SOM defines a mapping from the input data space

$$\mathbf{x} = [x_1, x_2, ..., x_m]^T$$

onto a regular two-dimensional array of nodes. The synaptic weight vector

$$\mathbf{w}_j = [w_{j1}, w_{j2}, ..., w_{jm}]^T, \; j = 1, 2, ..., l$$

of each neuron j has the same dimension as the input space; l is a total number of neurons.

Selecting the neuron with the largest inner product $\mathbf{w}_j^T \mathbf{x}$, is mathematically equivalent to minimizing the Euclidean distance between the input vectors $\mathbf{x}$ and $\mathbf{w}_j$. Thus, the winning neuron c is defined as:

$$c = arg\ min_j ||\mathbf{x} - \mathbf{w}_j||$$

Essentially this sums up the competition process among the neurons, where the best-matching node locates the center of a topological neighborhood [13].

Adaptive process

During the learning, those nodes that are topographically close to a certain distance will activate each other to learn from the same input. Using discrete-time formalism, weight vector at time t is written as $\mathbf{w}_j(t)$, and updated weight vector is defined as:

$$\mathbf{w}_j(t+1) = \mathbf{w}_j(t) + \mathbf{h}_{jc}(t)[\mathbf{x} - \mathbf{w}_j(t)]$$

where $\mathbf{h}_{jc}(t)$ is neighborhood kernel. For details see Ref. [15].

A SOM APPROACH TO B-TAGGING

Event data

In our SOM approach we feed SOM network with the same events and the same variables as used in the traditional track counting algorithm:

- Number of tracks in the jet cone
- Impact parameters and related IP significances for three leading tracks. (See Fig. 3.)

Figure 2 shows an example of CMS event used in this study.

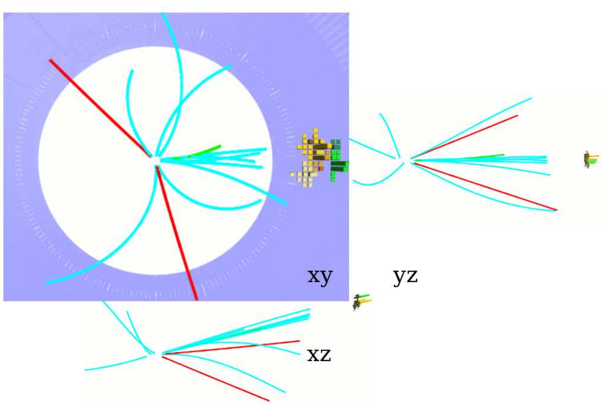

Figure 2: Successfully reconstructed jet can be identified as a b jets using lifetime based tagging algorithm, which relies on displaced secondary vertices. Two leptons leptons are presented by straight lines.

SOM configuration and Learning Vector Quantization LVQ

We used SOM_PAK [3] tool to analyze data created in CMS ORCA [1] simulation package, using full simulation with track and jet reconstruction.

In our hybrid approach, after the self-organizing feature map, the actual classification and feature map tuning (For details see Ref. [13].) was performed using a learning vector quantizer tool LVQ_PAK [3]. Details of this type of winner-takes-all algorithm are available elsewhere in the literature [15, 16].

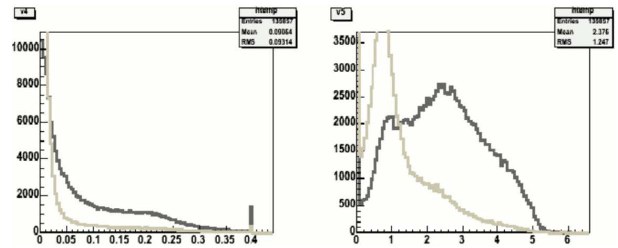

Figure 3: An example of variables used in the SOM teaching. A leading track IP distribution for a signal (left) and background events and a significance of a leading track IP (right).

We started learning with the optimized LVQ1 algorithm, which has a very fast convergence. Some performance improvement was achieved with subsequent use of LVQ2.1 algorithm.

RESULTS

The b tagging efficiency with SOM was found to be 73 % with 11 % mistagging rate. We were able to filter 45 % of the background events with 0.2% misclassification probability for the signal. Subsequent tuning of SOM with LVQ_PAK-tool provided 1.5 % improvement in tagging efficiency. In Figs. 4 and 5, which visualize SOM activation with test data, a clear separation to signal and background regions is seen. These results can be compared with typical counting algorithms performance 35 % efficiency with 1 % mistagging propability.

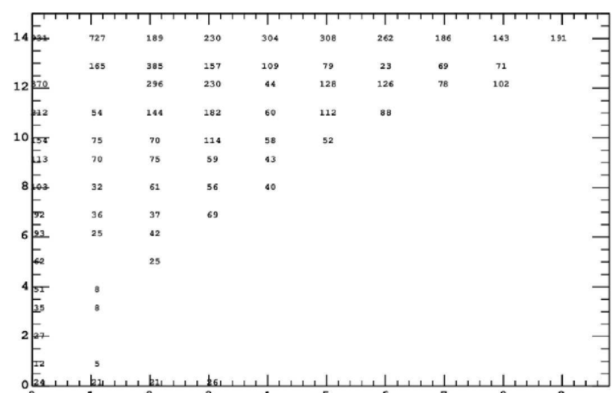

Figure 4: Number of signal events associated to winning node, while testing mapping of 15x15 node SOM performance.

CONCLUSION

We have shown that unsupervised classification can be utilized successfully in b-tagging problems. In our study the self-organizing maps are able to separate the Higgs signal from the background, based on CMS Monte Carlo data.

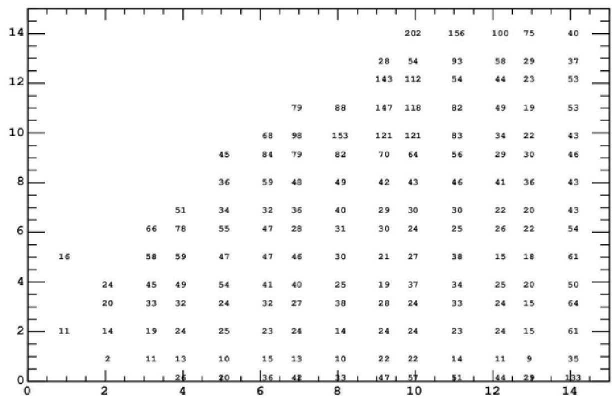

Figure 5: Number of background events associated to winning node, while testing mapping of 15x15 node SOM performance.

We consider HEP as a promising field where the potential of self-organizing networks has not yet been fully exploited.

Since the optimization of free parameters ϵ and δ and their time dependence can be very time consuming, our future work will consist of inclusion of growing SOM using a gravitational algorithm [11]. We are also planning to study very promising AdaBoost-algorithm as an alternative to artificial neural networks for the Higgs boson identification [17].

REFERENCES

[1] G. Segneri and F. Palla, “Lifetime based b-tagging with CMS”, CMS NOTE 2002/046, November 26, 2002.

[2] A. Heikkinen and S. Lehti, “Tagging b jets associated with heavy neutral MSSM Higgs bosons”. NIM A 559 (2006) 195–198.

[3] T. Kohonen *et. al.*, SOM_PAK: The Self-Organizing Map Program Package. Technical Report A31 *and* T. Kohonen *et. al.*, LVQ_PAK: The Learning Vector Quantization Program Package. Technical Report A30, Helsinki University of Technology, Laboratory of Computer and Information Science, FIN-02150 Espoo, Finland, 1996.

[4] C. Weiser, “A Combined Secondary Vertex Based B-Tagging Algorithm in CMS”, CMS NOTE 2006/014, January 25, 2006.

[5] J. Zimmermann and C. Kiesling, “Statistical learning methods in high-energy and astrophysics analysis”, NIM A 534 (2004) 204–210.

[6] K. Hultqvist *et al.*, “Using a neural network in the search for the Higgs boson”, NIM A 364 (1995) 193–200.

[7] I. Iashvili and A. Kharchilava, CMS TN-1996/100.

[8] M. Mjahed, Nuclear Physics B 140 (2005) 799–801.

[9] A. De Angelis *et al.*, “Self-organizing networks for classification: developing applications to science analysis for astroparticle physic”, Physics A 338 (2004) 50-53.

[10] K. -H. Becks *et al.*, “Separation of hadronic W-decays from QCD-background with self-organizing maps”, NIM A 426 (1999) 599–604.

[11] J. S. Lange, “Transputer self-organizing map algorithm for beam background rejection at the BELLE silicon vertex detector”, NIM A 420 (1999) 288–309.

[12] M. J. Lang, “Application of a Kohonen network classifier in TeV γ -ray astronomy”, J. Phys. G: Nucl. Part. Phys. 24 (1998) 2279–2287.

[13] S. Haykin, “Neural Networks”, Prentice-Hall (1999).

[14] T. Kohonen, “Self-Organizing Maps”, Springer-Verlag, Heidelberg (1995).

[15] J. Zurada, “Introduction to Artificial Neural Systems”, Jaico Publishing House, Mumbai (2003).

[16] M. Biehl *et al.*, “Learning vector quantization: The dynamics of winner-takes-all algorithms”, Neurocomputing 69 (2006) 660–670.

[17] B. Roe *et al.*, “Boosted decision trees as an alternative to artificial neural networks for particle identification”, NIM A 543 (2005) 577–584.