Discrimination of dimuon production events in neutrino interactions in the context of the MINERνA experiment

Sebastián Sánchez, Javier Rengifo, Nhell Cerna, y Marvin Ascencio

High Energy Group
Pontifícia Universidad Católica del Perú-PUCP

Lima - Perú
4th July 2016

Outline

Abstract

In the present work, we aim to optimize the discrimination of dimuon production events in neutrino - nuclei interactions in the context of the MINERvA experiment For this type of events, the main background are general dilepton production events. Our dataset was simulated with the GENIE Neutrino MC v2.8.6, which will also provide us with a simplified parameterization of the MINERvA experiment. The signal discrimination is done over experimentally reconstructable variables and optimized with TMVA [1] package v4.2.0. We finally discuss on the significance of our results for the search of new physics and nuclear structure.

Introduction

The production of lepton pairs induced by neutrino scattering in a Coulumb field of a target nucleus is called neutrino trident production and it offers further possibilities to study neutrino-nuclei interactions. It is described by the general reaction:

$$\nu_\ell(\bar{\nu}_\ell) + N \to \nu_\ell(\bar{\nu}_\ell) + \ell^+ + \ell^- + N \tag{1}$$

Where N is a nucleus in this $\ell = \mu$ because is most interest of experimental reasons.

Introduction

In standard model reaction can proceed via two interfering channels: charged (W) and neutral (Z) boson exchange (see Fig. 1).

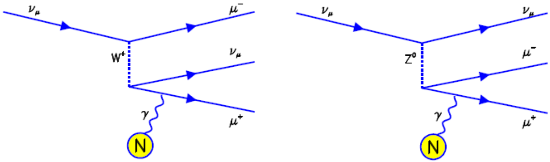


Figure: Feynman diagrams for neutrino trident production [2].

Introduction

A broader class of events is that of (visible) dilepton production from neutrino-nuclei interactions. These events usually include short lived meson decays or pair production from high energy photons. They constitute the major source of background for neutrino trident production. Considering also that the trident cross section is relatively small, $\sigma_{trident} \sim 10^{-6}\sigma_{CC}$ [3], it is imperative to have reliable methods for discriminating the signal (dimuon events) from background (dilepton events). This makes Multivariate Analysis (MVA) Methods attarctive as a way to extract as much information as possible from the tipically small samples of dilepton events.

Methodology

Data collection and Data Processing

The sample data files, first we testing with a little events called $gntp.NuMI_C12_dilepton_inclusive.root$. Secondly, we testing with data called $ultralisk_until7.gst.root$

We show a brief resume about the methods of TMVA [1] that we use in this work for the analysis of the data. Also we extract some definitions from TMVA web: http://tmva.sourceforge.net/

GENIE

```
|------------------------------------------------------------------------------------------------------------------------------|
|GENIE GHEP Event Record [print level:   3]                                                                                    |
|------------------------------------------------------------------------------------------------------------------------------|
| Idx |          Name | Ist |          PDG |   Mother  | Daughter  |        Px |        Py |        Pz |         E |          m |
|------------------------------------------------------------------------------------------------------------------------------|
|   0 |         nu_mu |   0 |           14 |  -1 |  -1 |   4 |   4 |     0.000 |     0.000 |     9.076 |     9.076 |      0.000 |
|   1 |           C12 |   0 |   1000060120 |  -1 |  -1 |   2 |   3 |     0.000 |     0.000 |     0.000 |    11.179 |     11.179 |
|   2 |        proton |  11 |         2212 |   1 |  -1 |   5 |   5 |    -0.039 |     0.057 |     0.016 |     0.924 |    **0.938 | M = 0.921
|   3 |           B11 |   2 |   1000050110 |   1 |  -1 |  24 |  24 |     0.039 |    -0.057 |    -0.016 |    10.255 |     10.255 |
|   4 |         nu_mu |   1 |           14 |   0 |  -1 |  -1 |  -1 |    -0.670 |    -0.084 |     1.194 |     1.372 |      0.000 | P = (0.488,0.061,-0.871)
|   5 |      HadrSyst |  12 |   2000000001 |   2 |  -1 |   6 |   7 |     0.631 |     0.140 |     7.897 |     8.627 |    **0.000 | M = 3.414
|   6 |             u |  12 |            2 |   5 |  -1 |   8 |   8 |     0.625 |     0.139 |     7.819 |     7.852 |      0.330 |
|   7 |          ud_1 |  12 |         2103 |   5 |  -1 |   8 |   8 |     0.006 |     0.001 |     0.078 |     0.775 |      0.771 |
|   8 |        string |  12 |           92 |   6 |  -1 |   9 |  12 |     0.631 |     0.140 |     7.897 |     8.627 |    **0.000 | M = 3.414
|   9 |           pi+ |  14 |          211 |   8 |  -1 |  17 |  17 |     0.127 |     0.024 |     0.058 |     0.199 |      0.140 | FSI = 1
|  10 |           pi- |  14 |         -211 |   8 |  -1 |  18 |  18 |     0.152 |     0.461 |     5.735 |     5.758 |      0.140 | FSI = 1
|  11 |        proton |  14 |         2212 |   8 |  -1 |  19 |  19 |     0.037 |    -0.190 |     1.524 |     1.800 |      0.938 | FSI = 1
|  12 |           eta |  12 |          221 |   8 |  -1 |  13 |  16 |     0.315 |    -0.154 |     0.580 |     0.871 |      0.547 |
|  13 |           pi- |  14 |         -211 |  12 |  -1 |  20 |  20 |     0.244 |    -0.179 |     0.295 |     0.445 |      0.140 | FSI = 1
|  14 |           pi+ |  14 |          211 |  12 |  -1 |  21 |  21 |     0.089 |    -0.049 |     0.285 |     0.334 |      0.140 | FSI = 5
|  15 |            e+ |   1 |          -11 |  12 |  -1 |  -1 |  -1 |    -0.018 |     0.077 |    -0.013 |     0.080 |      0.001 |
|  16 |            e- |   1 |           11 |  12 |  -1 |  -1 |  -1 |    -0.001 |    -0.003 |     0.012 |     0.013 |      0.001 |
|  17 |           pi+ |   1 |          211 |   9 |  -1 |  -1 |  -1 |     0.127 |     0.024 |     0.058 |     0.199 |      0.140 |
|  18 |           pi- |   1 |         -211 |  10 |  -1 |  -1 |  -1 |     0.152 |     0.461 |     5.735 |     5.758 |      0.140 |
|  19 |        proton |   1 |         2212 |  11 |  -1 |  -1 |  -1 |     0.037 |    -0.190 |     1.524 |     1.800 |      0.938 |
|  20 |           pi- |   1 |         -211 |  13 |  -1 |  -1 |  -1 |     0.244 |    -0.179 |     0.295 |     0.445 |      0.140 |
|  21 |      HadrClus |  16 |   2000000300 |  14 |  -1 |  22 |  23 |     0.089 |    -0.049 |     0.285 |     0.334 |    **0.000 | M = 0.140
|  22 |        proton |   1 |         2212 |  21 |  -1 |  -1 |  -1 |    -0.186 |     0.303 |     0.541 |     1.140 |      0.938 |
|  23 |       neutron |   1 |         2112 |  21 |  -1 |  -1 |  -1 |     0.276 |    -0.352 |    -0.256 |     1.072 |      0.940 |
|  24 |      HadrBlob |  15 |   2000000002 |   3 |  -1 |  -1 |  -1 |     0.039 |    -0.057 |    -0.016 |     8.378 |    **0.000 | M = 8.377
|------------------------------------------------------------------------------------------------------------------------------|
|       Fin-Init:                                                 |     0.000 |     0.000 |     0.000 |     0.000 |            |
|------------------------------------------------------------------------------------------------------------------------------|
|       Vertex:          nu_mu @ (x =     0.00000 m, y =     0.00000 m, z =     0.00000 m, t =              0 s)               |
|------------------------------------------------------------------------------------------------------------------------------|
| Err flag [bits:15->0] : 0000000000000000    |  1st set:                                                               none |
| Err mask [bits:15->0] : 1111111111111111    |  Is unphysical:    NO |   Accepted:   YES                                  |
|------------------------------------------------------------------------------------------------------------------------------|
| sig(Ev) =       5.1389e-38 cm^2  | d2sig(x,y;E)/dxdy =       1.1491e-37 cm^2      | Weight =                           1 |
|------------------------------------------------------------------------------------------------------------------------------|
```

GENIE

```
GENIE Interaction Summary
------------------------------------------------------------------------------------------------------------
[-] [Init-State]
 |--> probe        : PDG-code = 14 (nu_mu)
 |--> nucl. target : Z = 6, A = 12, PDG-Code = 1000060120 (C12)
 |--> hit nucleon  : PDC-Code = 2212 (proton)
 |--> hit quark    : PDC-Code = 2 (u) [valence]
 |--> probe 4P     : (E =      9.07572, Px =            0, Py =            0, Pz =      9.07572)
 |--> target 4P    : (E =       11.179, Px =            0, Py =            0, Pz =            0)
 |--> nucleon 4P   : (E =     0.923588, Px =   -0.0387719, Py =    0.0565415, Pz =    0.0156137)
[-] [Process-Info]
 |--> Interaction : Weak[NC]
 |--> Scattering  : DIS
[-] [Kinematics]
 |-->  *Running* Hadronic invariant mass W = 3.41351
 |--> *Selected* Bjorken x = 0.229753
 |--> *Selected* Inelasticity y = 0.851075
 |--> *Selected* Momentum transfer Q2 (>0) = 3.22265
 |--> *Selected* Hadronic invariant mass W = 3.41351
[-] [Exclusive Process Info]
 |--> charm prod.  : false
 |--> f/s nucleons : N(p) = 0 N(n) = 0
 |--> f/s pions    : N(pi^0) = 0 N(pi^+) = 0 N(pi^-) = 0
 |--> resonance    : [not set]
```

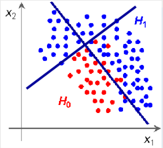
Methodology

Rectangular cut optimisation

Simplest method: cut in rectangular volume using

$$x_{\mathrm{cut}}(i_{\mathrm{event}}) \in \{0,1\} = \prod_{v \in \{\mathrm{variables}\}} \left(x_v(i_{\mathrm{event}}) \subset \left[x_{v,\mathrm{min}}, x_{v,\mathrm{max}} \right] \right)$$

Cuts usually benefit from prior decorrelation of cut variables



Discussion

In this section we show plots that describe the most important features about the output that we obtained using TMVA. From figure 2, we can see that the lepton pair invariant mass is the best input variable to separate signal and background, because is posible fix a simple cut ans separate the signal with a little contamination.

Discussion

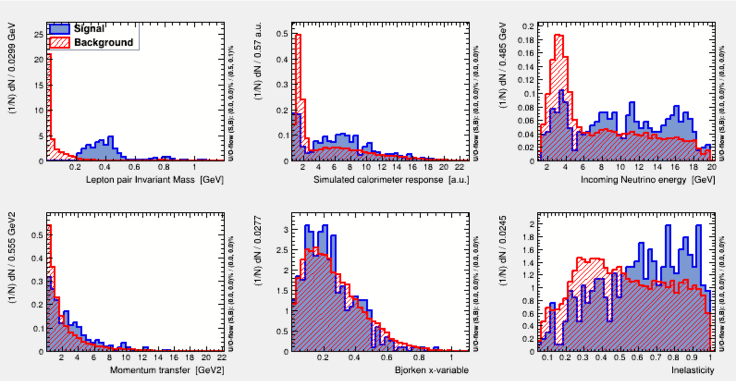


Figure: Plots of Input variables (training sample) show histograms of signal in blue and bakground in red

Discussion

Signal

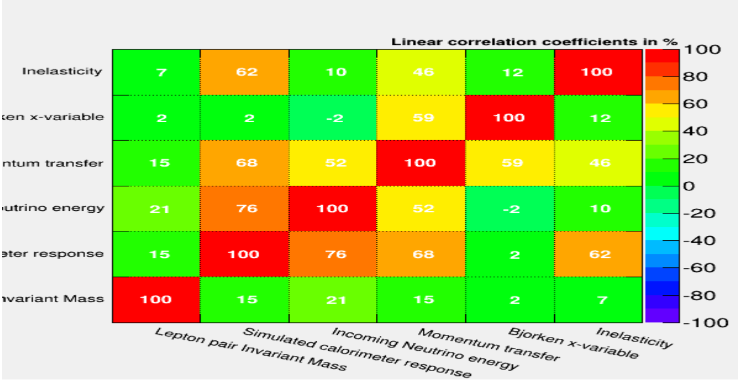


Figure: Input variable linear correlation coefficients

Discussion

Background

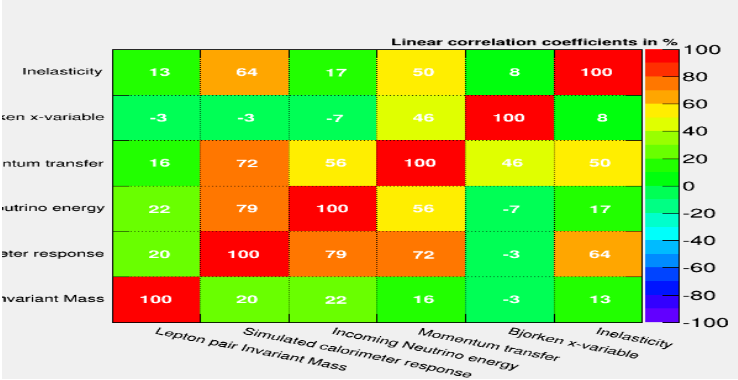


Figure: Input variable linear correlation coefficients

Discussion

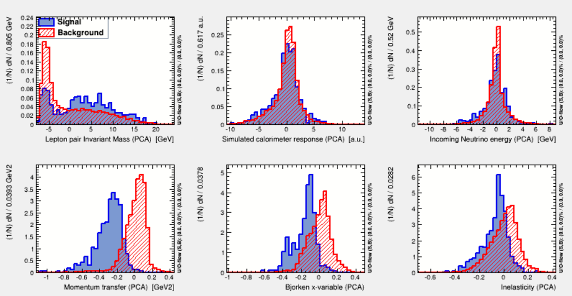


Figure: Plots of Input variables (training sample) show histograms of signal in blue and bakground in red with PCA transformed

Discussion

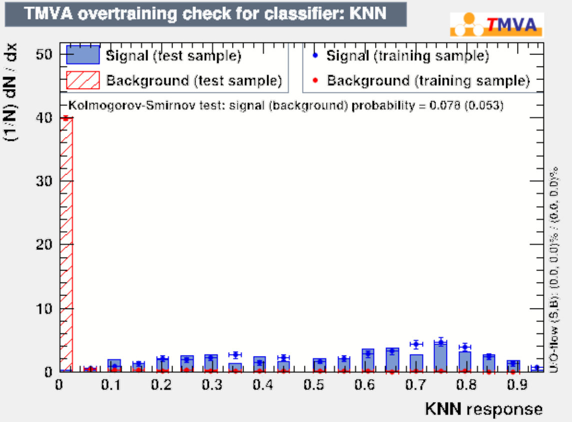


Figure: TMVA overtraining check for classifier: KNN

Discussion

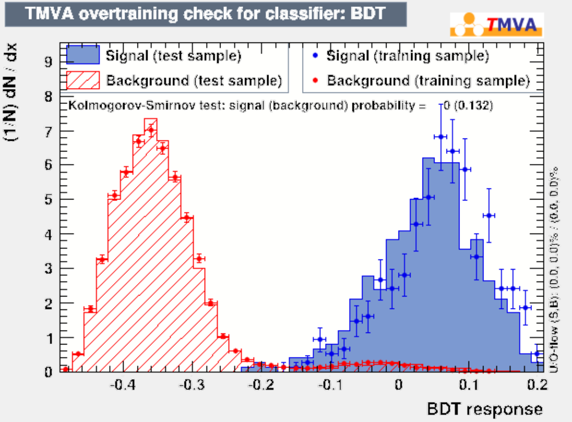


Figure: TMVA overtraining check for classifier: BDT

Conclusion

In the present work we have briefly explored MVA methods for signal to background discrimination. It was found that non linear methos, like BDT, kNN and MLPBNN offer the possibility to disentangle the signal better than most linear methods, and far better than possible with the application of a cut over a single variable. Specifically, from the input variable's distribtuions alone, it may seem obvious that the only useful variable was the lepton pair invariant mass. Nevertheless, the MultiVariate Analysis and Optimization takes advantage of the additional not automatically evident information to greatly improve the separation.

Conclusion

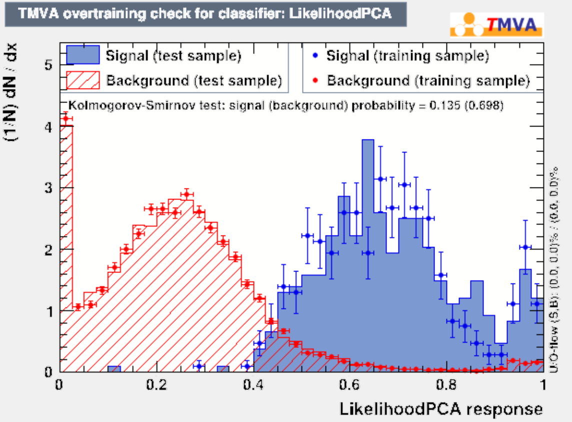


Figure: TMVA overtraining check for classifier: KNN

Conclusion

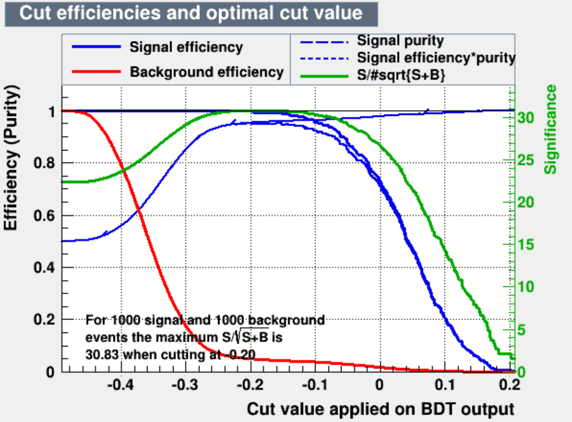


Figure: Cut optimization for classifier: BDT

Conclusion

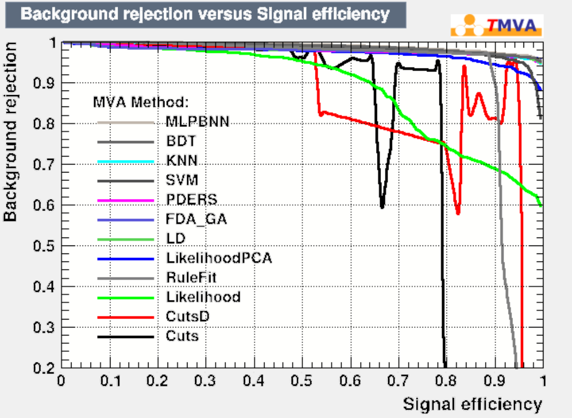


Figure: Background rejection versus signal efficiency

Reference

- HOECKER, A. et al. TMVA: Toolkit for Multivariate Data Analysis. *PoS*, ACAT, p. 040, 2007.
- ADAMS, T. et al. Neutrino trident production from NuTeV. In: *High-energy physics. Proceedings, 29th International Conference, ICHEP'98, Vancouver, Canada, July 23-29, 1998. Vol. 1, 2.* [S.l.: s.n.], 1998.
- BELUSEVIC, R.; SMITH, J. W - Z Interference in Neutrino - Nucleus Scattering. *Phys. Rev.*, D37, p. 2419, 1988.