Subject: TMVA - weight(s)
Posted by Elisabetta Prencipe (2) on Sat, 07 Mar 2015 11:09:48 GMT
View Forum Message <> Reply to Message

Dear multi-variate-analysis experts,

recently I did some studies to train a NN and Fischer discriminant, using a sample of signal and a sample of background, not properly normalized to the same equivalent luminosity. Basically the standard standalone tool works fine. I downloaded it from here: http://tmva.sourceforge.net

In the test-macro they provide (here it is attached), there is the possibility to assign different weights to signal and background sample. I set up the signal weight to 1, and the bkg weight to 300 (in my particular case). However, it looks that the output plot never changes, depending on different weights applied to the background root-tree (I tried with w=50, 300, 300); this is not what I would expect. I attac here the KNN output plot.

The part of the code where one can set up the weights, for example, is here:

```
// --- Register the training and test trees

TTree *signal     = (TTree*)input->Get("TreeS");
TTree *background = (TTree*)input->Get("TreeB");


// global event weights per tree (see below for setting event-wise weights)
Double_t signalWeight     = 1.0;
Double_t backgroundWeight = 50.0;


// You can add an arbitrary number of signal or background trees
factory->AddSignalTree    ( signal,     signalWeight     );
factory->AddBackgroundTree( background, backgroundWeight );
```

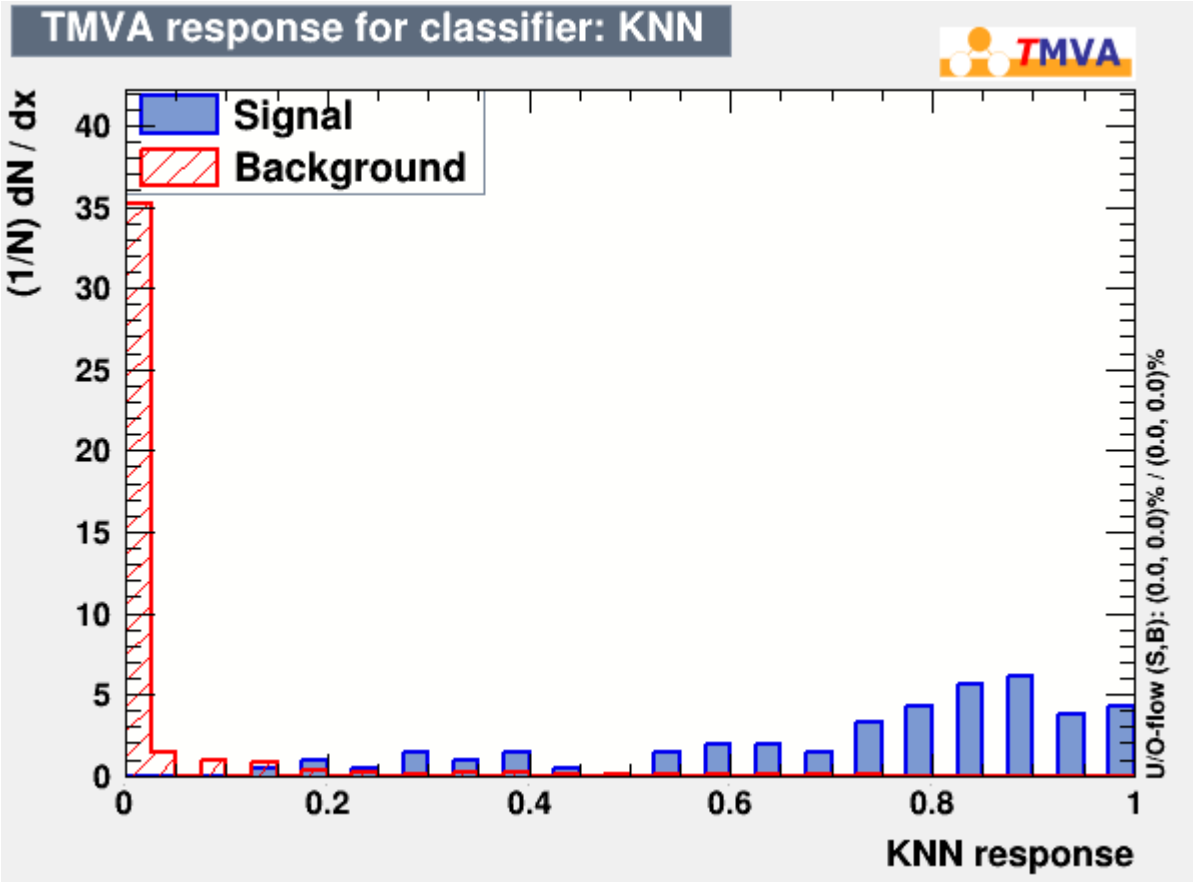
What am I doing wrong here, in your opinion?

Thank you in advance for any feedback,

Elisabetta

File Attachments

1) TMVAClassification.C, downloaded 798 times
2) plotKNN.gif, downloaded 1323 times



Subject: Re: TMVA - weight(s)
Posted by StefanoSpataro on Sat, 07 Mar 2015 12:12:12 GMT
View Forum Message <> Reply to Message

I'm not surprised, since the plot you are showing has normalized histograms. Try to check what is happening in the ROC curve.

Subject: Re: TMVA - weight(s)
Posted by Elisabetta Prencipe (2) on Sat, 07 Mar 2015 13:27:26 GMT
View Forum Message <> Reply to Message

Hi Stefano,

same story: ROC curve does not change if the bkg weight is 1, 50 or 500, for example. This is nasty....
I have an old tool for the Fischer discriminat only. The possibility to add a weight was not, at that time. So, if I trained my F-discriminant with different size of bkg samples, the efficiency curve and the discriminant shape, looked different. I am confused on how TMVA works, honestly.

Here are two possibilities:

a) something is wrong with this tool;

b) the tool is ok, I do not understand how it works.

I am in favor of b)....

Any suggestion?

Elisabetta

Subject: Re: TMVA - weight(s)
Posted by Elisabetta Prencipe (2) on Sat, 07 Mar 2015 13:34:35 GMT
View Forum Message <> Reply to Message

Ah,
 I think I found what shall I look to get differences.
 The curve called "classifier cut efficiencies" shows the efficency of sig and bkg, and gives you a rat-window where you can change the scaling factor S/B. And they vary depending on the ratio that you suggest as input.

Good! One problem less!

Elisabetta
