3 TMVA exercise – Multivariate Classification

This exercise uses the same data as in part A, exercise 2: Gaussian distributions with added correlations. The macro is still `TMVAClassification_correlatedGauss.C`, using the file `data_correlatedGauss.root`.

3.1 Comparison with Boosted Decision Trees and Neural Networks

Two more methods are brought into comparison: Boosted Decision Trees (BDT) and Neural Networks (Multilayer Perceptron or MLP). They can be added to a single macro execution by invoking

```
$ root TMVAClassification_correlatedGauss.C\(\"CutsGA,BDT,MLP\"\)
```

1. Train the classifiers and look at the training and testing errors. Do you suspect any overtraining? Why or why not?

2. Study both the Decision Tree and Neural Network architectures. Do they make sense to you?

3. Pick the *Receiver-Operation-Characteristics* (ROC) curve form the GUI. All methods that are used in the invocation are displayed (here: CutsGA, MLP and BDT).

   (a) What would be the *ideal* ROC curve?
   (b) Which of the three methods shows better performance?
   (c) Can you explain in words why some methods might give better results that others?

4. Take a look at the parameters used to train the BDT and MLP. Try changing them one at a time in the macro and look at the effects on the ROC curve.