Exercise 9: Multivariate analysis

N. Berger (`nberger@physi.uni-heidelberg.de`)

5.12.2011

Please send your solutions to `nberger@physi.uni-heidelberg.de` until 12. 12. 2011, 12:00. Put your answers in an email (subject line *SMIPP:Exercise09*).

1. **TMVA** Root comes with an extensive package for multivariate analysis called TMVA (`tmva.sourceforge.net`. On the CIP-Pool machines, TMVA is installed. To test and explore it, copy the directory `/opt/root-5.28/tmva/test` into a location that is writable by you. Change into the directory and run
   `> root -l TMVAClassification.C\(\"Fisher\"\)`.
   You should get a window with lots and lots of buttons, each of which either produces plots or another set of buttons producing even more plots for the test sample that comes with TMVA. Have a look at them and try to find out what they mean (there is extensive documentation on the TMVA website).

2. **TMVA II** Run `> root -l TMVAClassification.C`. This will train and run all classifiers available in TMVA and take about 15 minutes or so. At 90% efficiency, which classifier has the best purity?

3. **Bump hunting again** On the course website, there is a pure background and a pure signal file for samples similar to exercise 8. Modify `TMVAClassification.C` to use these samples (use the masses as spectator variables) and then apply it to the two trees from exercise 8. How easy do you find it to change little bits and pieces in a large package written by someone else? How well does your cut-based work from last time compare to TMVA?
   (Attach the modified `TMVAClassification.C` file and a few representative plots - especially a mass spectrum.)