Ideal τ-tagging with TMVA multivariate data-analysis toolkit *

A. Heikkinen, P. Kaitaniemi, V. Karimäki, R. Kinnunen,
M. J. Kortelainen, T. Lampén, S. Lehti, T. Lindén, and L. Wendland
Helsinki Institute of Physics, P.O. Box 64, FIN-00014 University of Helsinki, Finland. E-mail: aatos.heikkinen@cern.ch

We report on our experience of using ROOT package TMVA for multivariate data analysis, for a problem of τ-tagging. Using Monte Carlo data we investigate how in the ideal case τ tagging could be performed and Higgs signal separated from QCD background present in LHC experiments. Particularly we are interested of ideal tagging effciency achievable at 1% mistagging rate for various MVA classifiers.

I. INTRODUCTION

We have previously tested TMVA software in b-tagging for the search of MSSM Higgs bosons at the LHC [2].

II. MULTIVARIATE DATA ANALYSIS

A. TMVA methodology

III. DATA AND VARIABLES

A. Event generation

B. Ideal case vs. detector effects

C. Discriminating variables

D. IMPLEMENTATION

[code example]
[listing waht was coded, changed in TMVA]

IV. RESULTS

[Critics: TMVA bugs, is it already mature? Problems?]
[Praisal: what is now working compared to CHEP'07]

V. CONCLUSION AND DISCUSSION

TMVA

* Paper [1] in preparation for CHEP 2009, 21 - 27 March 2009 Prague, Czech Republic

1. References and notes

[1] A. Heikkinen et al. Ideal τ-tagging with tmva multivariate data-analysis toolkit. *Paper in preparation for International Conference on Computing in High Energy and Nuclear Physics (CHEP09), 21 - 27 March 2009 Prague, Czech Republic.* http://www.helsinki.fi/~miheikki/system/refs/heikkinen/, `ah09bProceedings.pdf, ah09bProceedings.tar.gz`.

[2] A. Heikkinen with T. Lampen et al. Testing TMVA software in b-tagging for the search of MSSM Higgs bosons at the LHC. International Conference on Computing in High Energy and Nuclear Physics (CHEP07), Journal of Physics: Conference Series 119 (2008) 032028, [doi:10.1088/1742-6596/119/3/032028].

APPENDIX A: WORKING NOTES

Suggested responsibility:

aatos Aatos: editor, NN classifiers;

- Pekka: git consulting, PROOF
- Sami: MC data,
- Lauri 1-prog physics
- Ritva:
- Tomas: Ametisti
- Tapio:
- Matti:a mechanism to work with variables
- Veikko:

1. HISTORY

- 081028 Project released in http://github.com/aatos/chep09tmva. Removed proceedings notes in the Appendix A to separate file `notes.tex`.
- 081021 Title and abstract focus improved after discussion in the group.
- 081014 First draft done after the idea to have TMVA paper at next CHEP was accepted in the group.

To be done:

- Template code for analysis using latest ROOT, and TMVA inside it.
- Revise title, abtract and paper structure including appendix.

2. Code repository

Source code for paper and TMVA scriphttp://github.com/aatos/chep09tmva
Guide http://ktown.kde.org/~zrusin/git/git-cheat-sheet-medium.png.

3. Current status of TMVA

For introduction browse, six talks from year 2008 http://tmva.sourceforge.net/talks.shtml.

- Current version is TMVA-v3.9.5 (2008 Aug. 9th).
- TMVA (http://tmva.cvs.sourceforge.net) is now released as ROOT package
  - ROOT version from 5-19-02a to 5-21-01-alice contains TMVA 3.9.4.
- In addition to many bug fixes:
  - Improved prepossessing
  - Pre-selection cuts on arrays. Previously used *TEventlists* (only event wise pass/fail) were replaced by *TreeFormulas* (sensitive to array position).
  - Plugin capability: custom multivariate classifier can now be plugged into the TMVA framework to benefit from TMVA's analysis and performance comparison tools.
  - For details see release notes http://tmva.cvs.sourceforge.net/*checkout*/tmva/TMVA/development/RELNOTES

4. Code

A code and data will be distributed using git and made abvailable at
http://www.helsinki.fi/~miheikki/system/refs/heikkinen/ah09bProceedings/code