TMVA New Features

Sergei V. Gleyzer
Omar Zapata Mesa
Lorenzo Moneta

Data Science at the LHC Workshop
Nov. 9, 2015
Outline

• New TMVA Features
• Data Loader
• Feature Selection
• RMVA
• PyMVA
New Features

• Improved design that allows
  – Greater flexibility (modularization)
  – Feature Selection, Cross-Validation, new techniques for data storage and manipulation
  – Parallelization (OpenMP/MPI/cuda/TBB)
  – Integration with Python and R (PyMVA/RMVA)
Data Loader

• TMVA::DataLoader class allows greater flexibility in dealing with data
  – Connection of different features to different classifier methods: Useful for optimization
Feature Selection

• Based on the FAST stochastic wrapper algorithm
  – See previous talk and today’s tutorial for details
RMVA

• RMVA is a set of plugins for TMVA that allows the use of R’s classification and regression packages
### Evaluation results ranked by best signal efficiency and purity (area)

<table>
<thead>
<tr>
<th>MVA Method</th>
<th>Signal efficiency at bkg eff.(error):</th>
<th>Separation</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>@B=0.01     @B=0.10     @B=0.30   ROC-integ.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RSVM</td>
<td>0.323(08)   0.735(08)   0.924(04) 0.913</td>
<td>0.526</td>
<td>1.355</td>
</tr>
<tr>
<td>RMLP</td>
<td>0.286(08)   0.689(08)   0.899(05) 0.897</td>
<td>0.481</td>
<td>1.310</td>
</tr>
<tr>
<td>C50</td>
<td>0.080(00)   0.671(08)   0.878(05) 0.881</td>
<td>0.462</td>
<td>1.253</td>
</tr>
<tr>
<td>RXGB</td>
<td>0.233(07)   0.643(08)   0.867(06) 0.875</td>
<td>0.434</td>
<td>1.194</td>
</tr>
</tbody>
</table>
PyMVA

- PyMVA is a set of plugins for TMVA based on python api that allows use of python based ML methods
**PyMVA**

**Evaluation results ranked by best signal efficiency and purity (area)**

<table>
<thead>
<tr>
<th>MVA Method</th>
<th>Signal efficiency at bkg eff.(error):</th>
<th>Separation</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>@B=0.01 @B=0.10 @B=0.30 ROC-integ.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PyGTB</td>
<td>0.343(08) 0.751(07) 0.924(04) 0.914</td>
<td>0.539</td>
<td>1.514</td>
</tr>
<tr>
<td>PyAdaBoost</td>
<td>0.331(08) 0.741(07) 0.918(05) 0.911</td>
<td>0.761</td>
<td>0.943</td>
</tr>
<tr>
<td>PyRandomForest</td>
<td>0.245(07) 0.702(08) 0.905(05) 0.898</td>
<td>0.497</td>
<td>1.375</td>
</tr>
</tbody>
</table>
Tutorials

Try the new features tutorial today and the interfaces tutorial on Thursday

More features coming soon.
More Information

Websites: [http://root.cern.ch](http://root.cern.ch)  
[http://ooproject.org](http://ooproject.org)