TEVATRON Higgs Analysis

First results - use of the Optimal Brain Surgeon pruning algorithm

Marcin Wolter

Friday, August 10, 2001

Simulated data used for analysis:

- signal: wh110.rwn, 4538 events,
- background: ttb.rwn, 1860 events.

Variables used in the analysis:

1. etb1 - b_1 transverse energy
2. etb2 - b_2 transverse energy
3. mjj - $M_{b\bar{b}}$
4. ht - H_T
5. ete - E_T^l
6. sph - S - sphericity
7. drb1b2 - $\Delta R(b_1,b_2)$

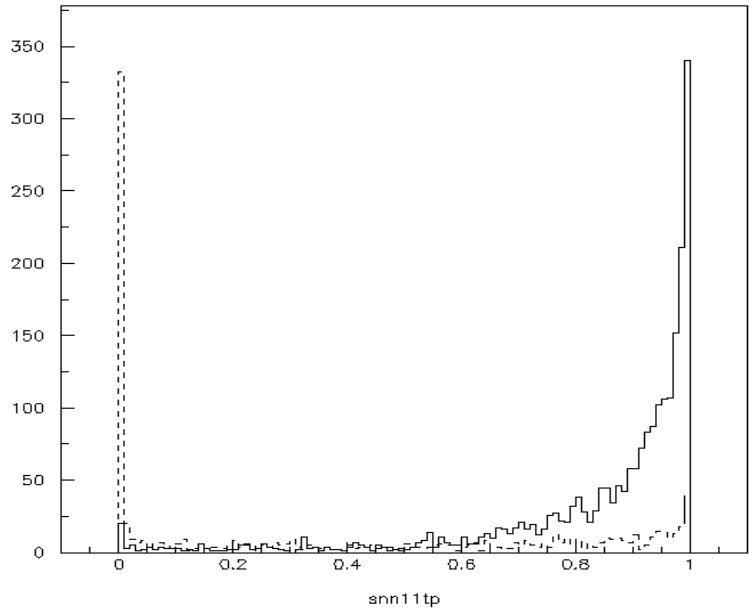
- snn11tp - output of the neural network used in the previous analysis.

F*or detailed description of the previous analysis and the variables used see "A Strategy for Discovering a Low-Mass Higgs Boson at the Tevatron", Fermilab-Pub-00/006.*

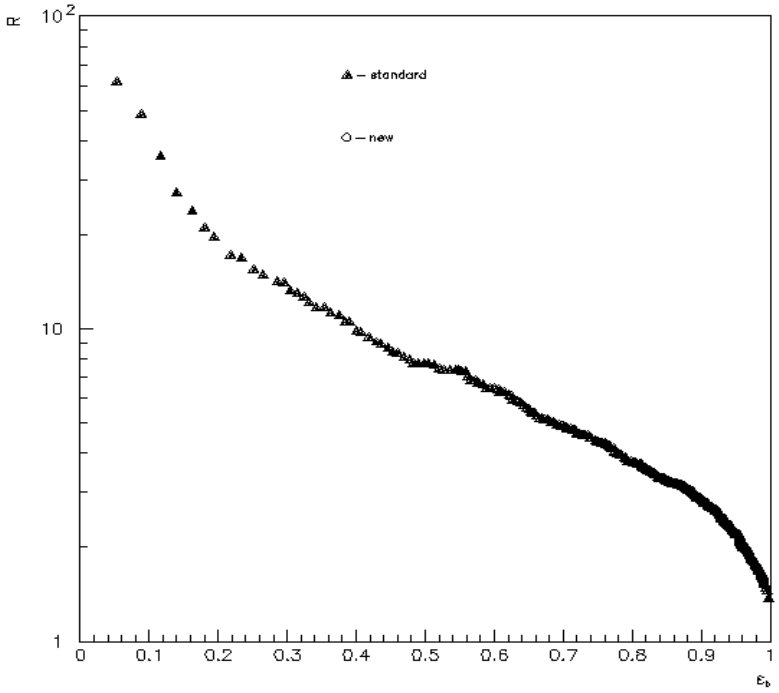
Results of the standard analysis.

The output of the network stored in the ntuple is used. This network is referred as "standard" in this text.

The following plots show the output weight distributions for signal and background and the performance shown as a dependence of the background rejection (Y-axis) as a function of the signal efficiency (X-axis).

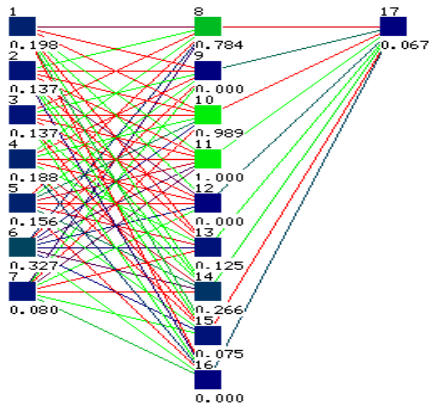


"Standard" network output weight distribution for signal (solid line) and background (dashed line).

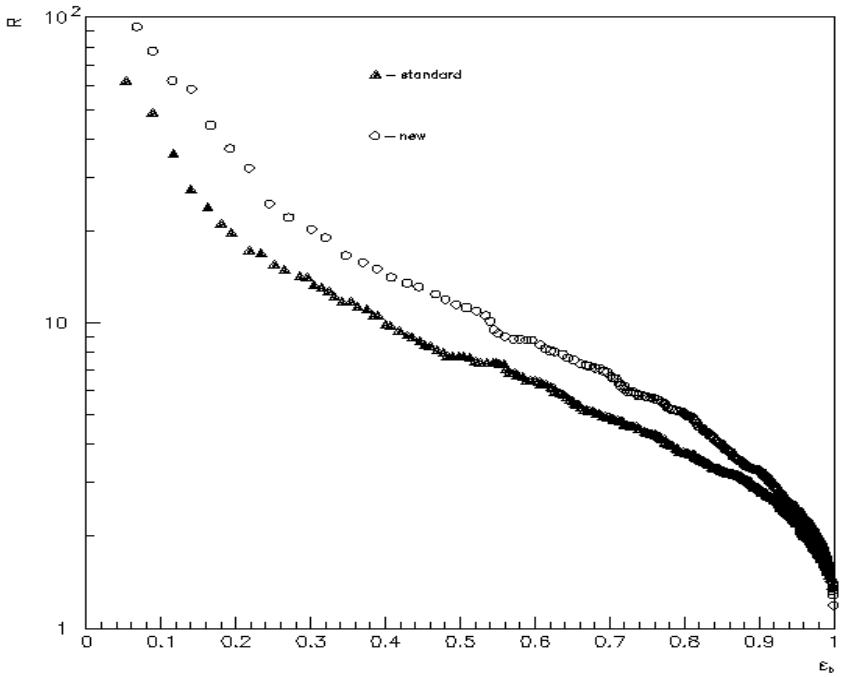


"Standard" network performance as a dependance of background rejection on signal efficiency.

Training feed-forward network.
The feed-forward network with 7 inputs and 9 nodes in the hidden layer is designed and trained. Data samples are divided into two sets (training and verifying). The first is used to train a network, the later to obtain the network performance.
The performance of the obtained network is compared with the "standard" one. The performance is better, than for the standard net.

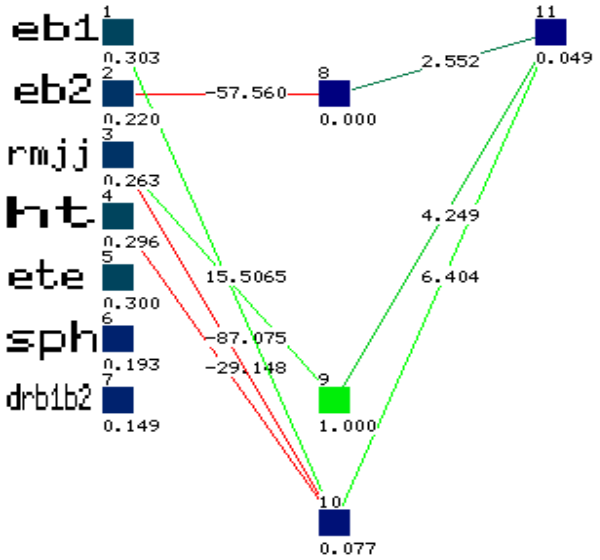


Network geometry.



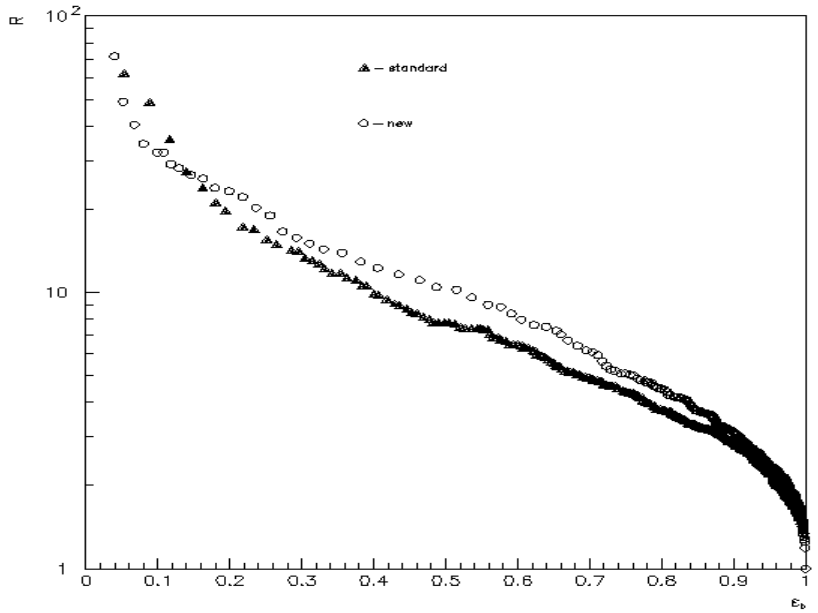
Performance of the feed-forward network with 7 units in the hidden layer and full connection (marked as "new") compared with the standard network.

Pruning algorithm
To the trained standard feed-forward network with 7 input units, 9 units in the hidden layer and one output unit a pruning algorithm (Optimal Brain Surgeon) is applied (see A.Zell "Simulation neuronaler Netze"). This algorithm removes the connections, which do have no significant effect on the output. After applying the pruning algorithm the number of links is significantly reduced, as well as three input units (ete, sph, drb1b2) are no longer used.
Smaller number of links means also faster learning and, even more important, frequently better generalization.

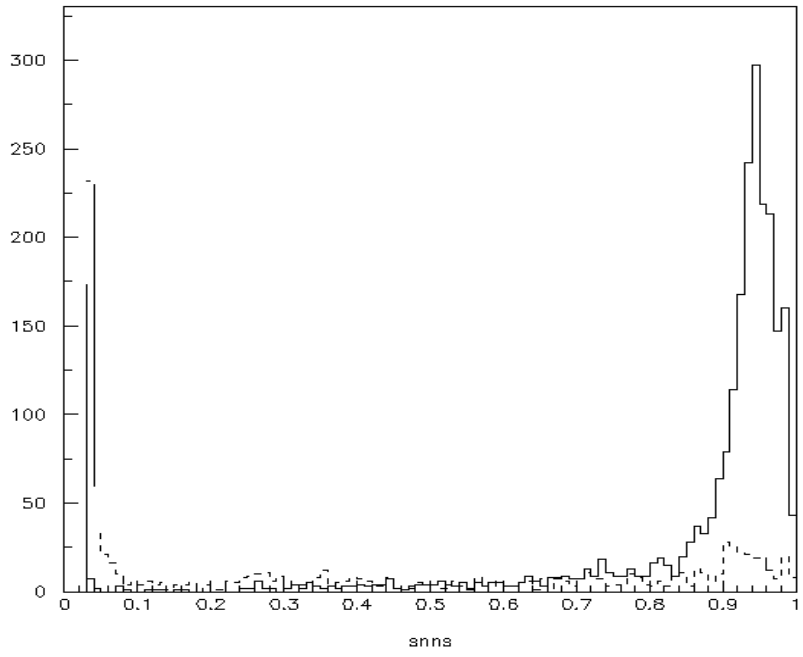


A network after pruning.

The obtained network, highly simplified, provides better performance than the "standard" network, and only slightly worse than the full network with nine units in the hidden layer described above. The plot below shows background rejection as a function of signal efficiency for a new network after pruning compared with the "standard" network. Probably the applying the pruning algorithm with softer cuts would give a network with a performance very similar to the full network.



Performance of the "pruned" network (marked as "new") compared to the "standard" network.



Output weight distribution for the network after rpunning. The solid line represents the distribution for signal events, while the dashed line for background events.

Conclusions

This was just a first look on the data. It was shown, that a good signal/background separation can be obtained also by using only four input variables and a highly simplified network.