BDT - NTrees

Giulia Maineri

Università degli studi di Milano

April 2022

Table of Contents

BookMethod

Training-BookMethod

Trees training was done with the following options:

- !H; in order not to print a method-specific help message
- !V; necessary to deactivate "verbose mode", which prints explanations of what's going on
- NTrees=50/150/450/850/1550; number of trees in the forest
- MinNodeSize=5%; minimum percentage of training events required in a leaf node
- MaxDepth=3; maximum depth of the decision tree allowed

Training-BookMethod

- BoostType=AdaBoost; boosting type
- AdaBoostBeta=0.2; learning rate for AdaBoost algorithm
- SeparationType=GiniIndex; separation criterion for node splitting
- nCuts=20;number of grid points in variable range used in finding optimal cut in node splitting

Variable $p_T^{balance}$ was deleted due to previous training that showed its discriminating power was very low.

Trainings

50 trees

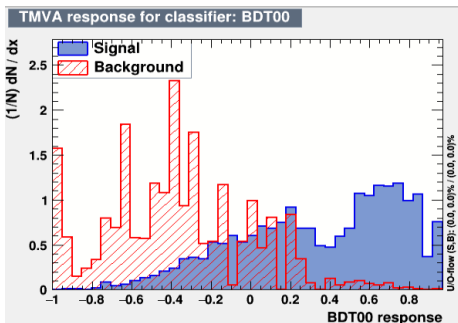# Classifier output distribution-Test and training sample

Figure 1: Output distribution

Figure 2: Overtraining control
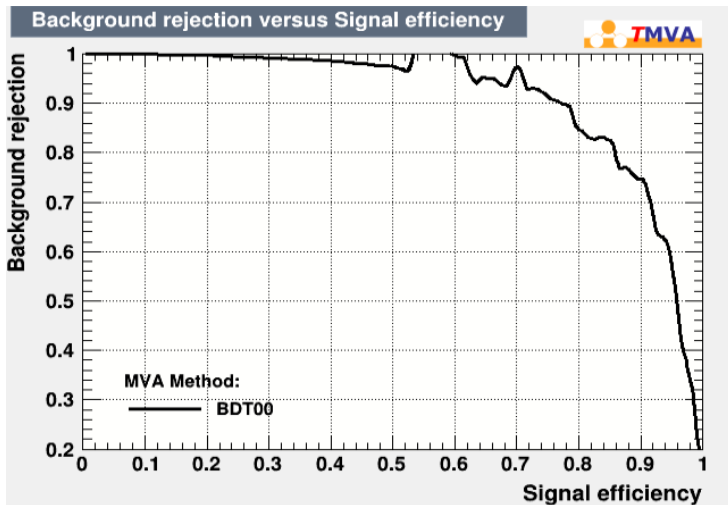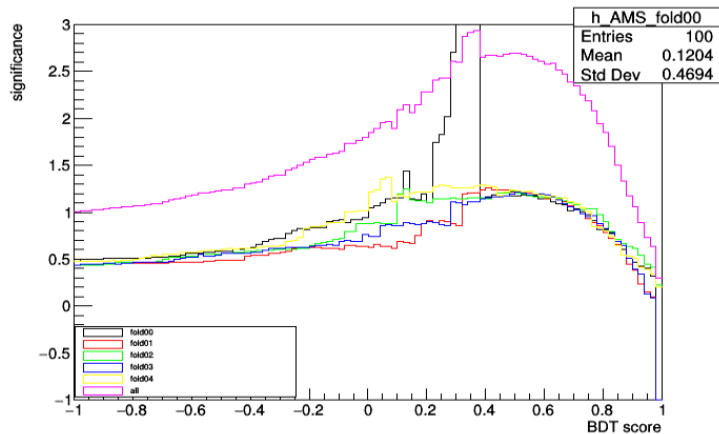
Figure 3: ROC curve

AMS
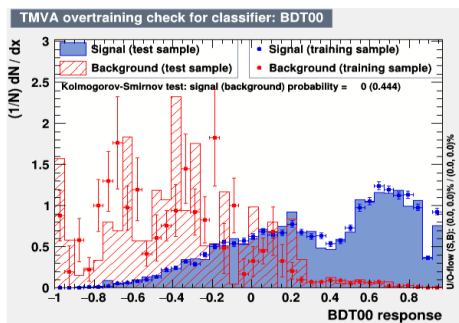
Figure 4: AMS

150 trees

Classifier output distribution-Test and training sample
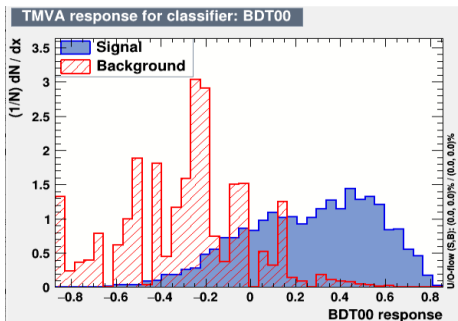
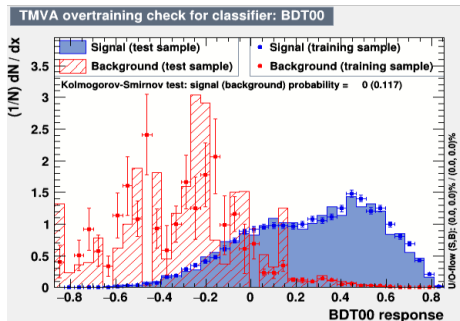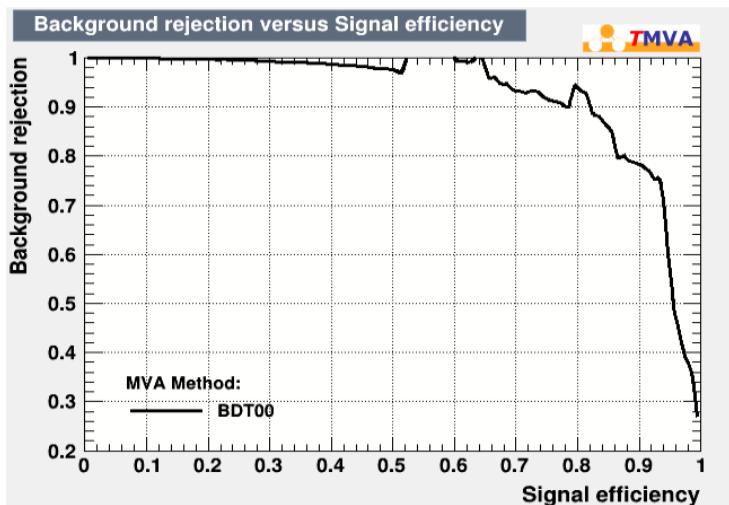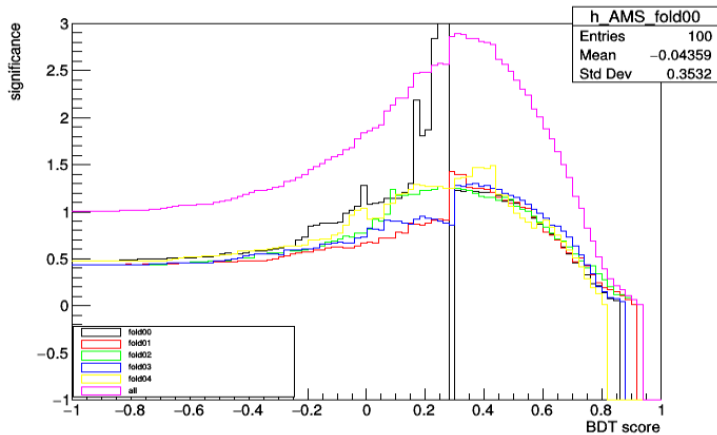Figure 5: Output distribution

Figure 6: Overtraining control

ROC curve

Figure 7: ROC curve

AMS

Figure 8: AMS

450 trees

Classifier output distribution-Test and training sample

Figure 9: Output distribution

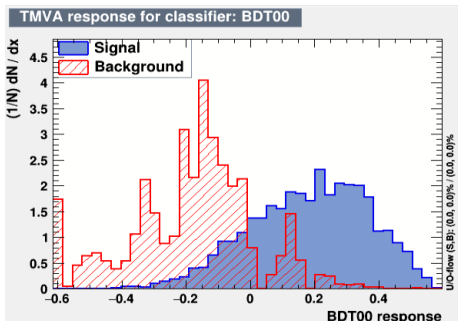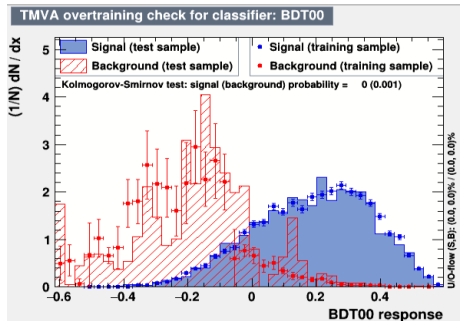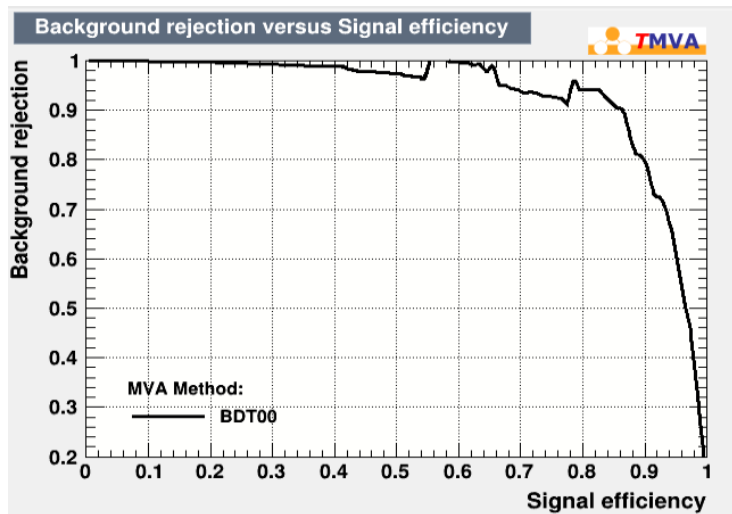Figure 10: Overtraining control

ROC curve

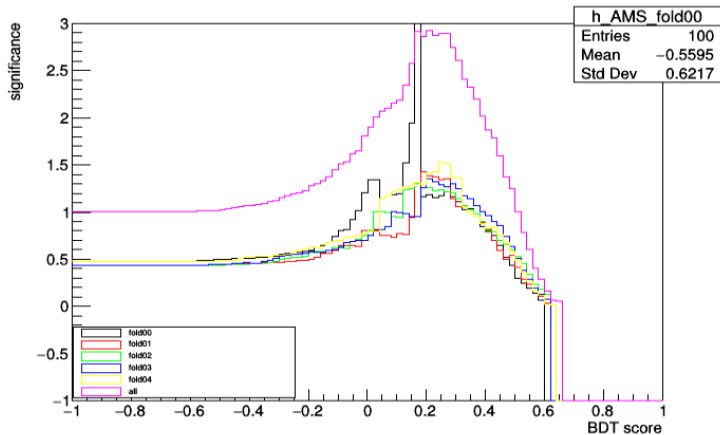Figure 11: ROC curve

AMS

Figure 12: AMS

850 trees

Classifier output distribution-Test and training sample
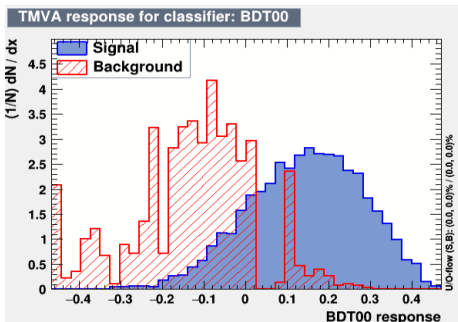
Figure 13: Output distribution
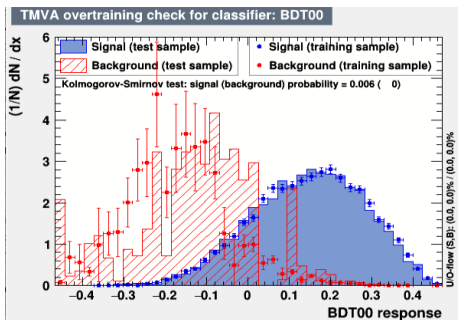

Figure 14: Overtraining control
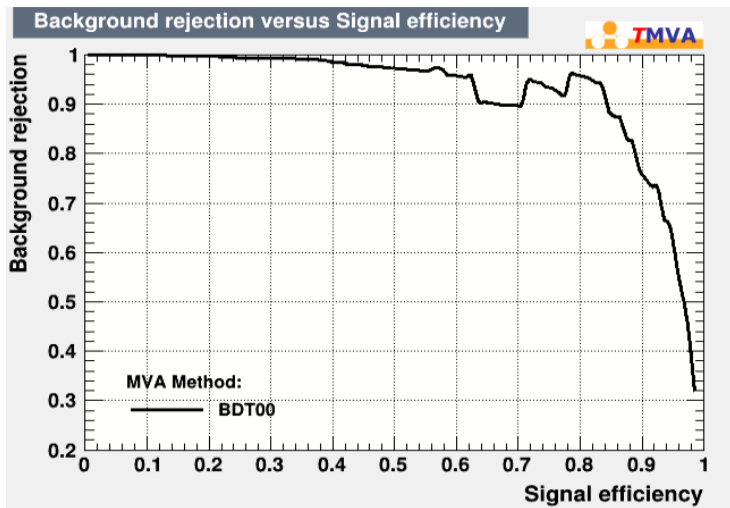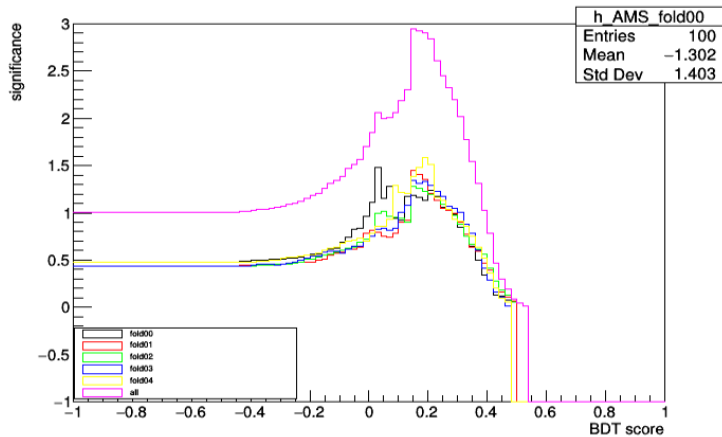
Figure 15: ROC curve

Figure 16: AMS

1550 trees

Classifier output distribution-Test and training sample
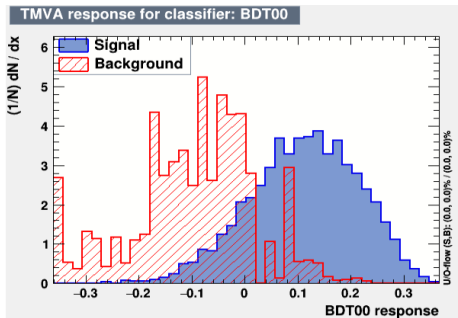
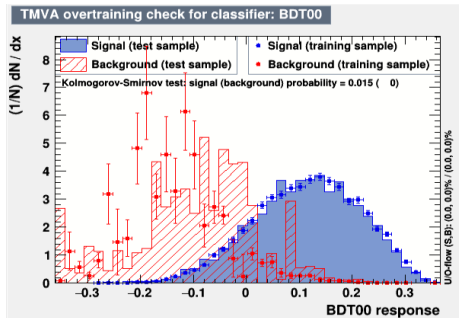Figure 17: Output distribution
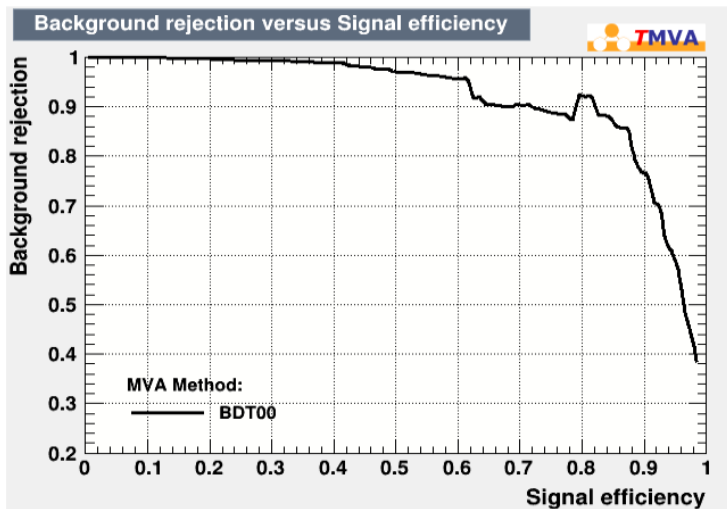
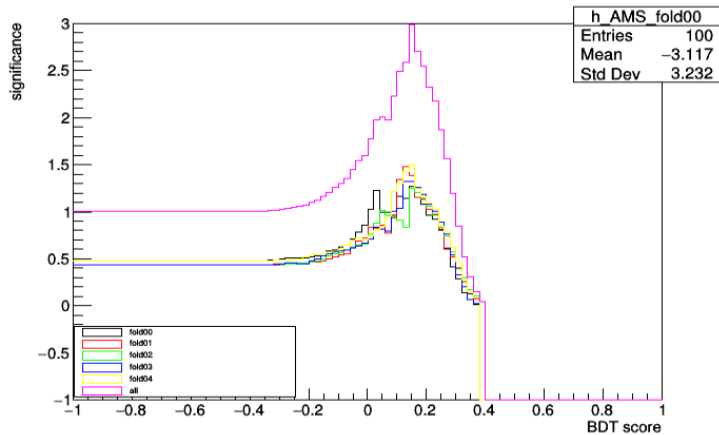

Figure 18: Overtraining control

Figure 19: ROC curve

AMS

Figure 20: AMS

Comparisons

Significance

NTrees	AMS
50	2.93757
150	2.89821
450	2.92140
850	2.94389
1550	2.98821

Table 1: Significances for different NTrees

Figure 21: Comparison

Comparison