Applications of SVMs in LEP data analysis

Stefan Soldner−Rembold

Fermilab

Run2aag Meeting, March 7, 2002

For a general introduction into SVM, see talk by H.Prosper

At LEP the performance of SVMs has been studied for the following problems:

M. Feindt, C. Haag (DELPHI note 99 −151):

− Separating B+ from other b decays

− Particle identification with dE/dx

SSR, P. Vannerem (OPAL), hep −ex/9905027

− Muon identification (P.Vannerem thesis)

− Charm tagging (hep −ex/9905027)

A non − separable problem: identify B+ in b decays

~20% signal in training and test samples

Variables used (see also DELPHI 99 −104):

1) inclusive vertex charge

2) inclusive vertex charge/error estimate

3) jet charge

4) flavour net output for same hemisphere

5) flavour net output for opposite hemisphere

6) weighted charge of leading fragmentation track

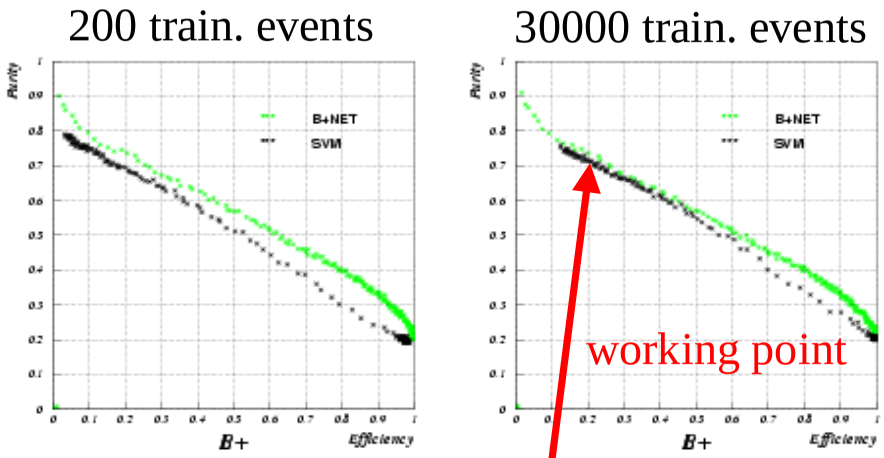


Figure 3: *a) Comparison between Network and SVM for B^+. On the left side a comparison for B^+ is shown, where the test sample for the SVM consisted of only 200 events. b) The right side shows the same comparison, but this time with 30000 events used for training. At the working point it achieves about the same performance as the NN.*

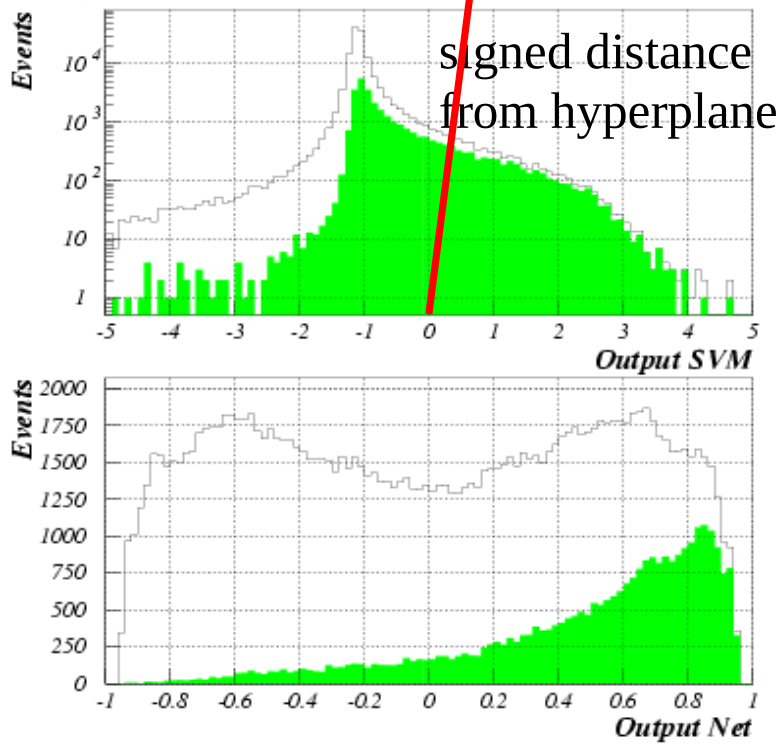
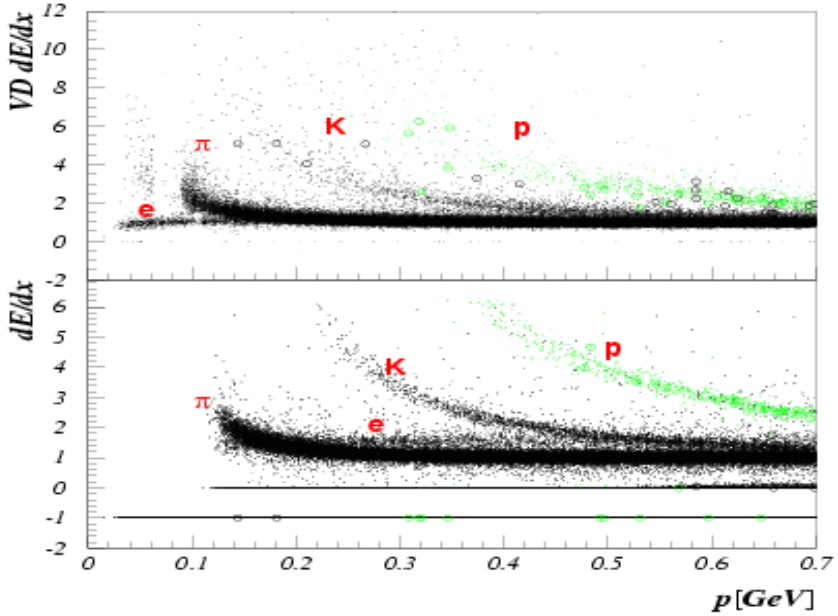


Figure 4: *Output of SVM (top) (note the logarithmic scale!) and of NN (bottom) where the signal (B^+) is shaded.*

- cut events with missing variables (opp. hemisphere tag)
- Gaussian kernel
- CPU time ~ N2 training with 200 and with 30000 events
- SVM performance at working point comparable to ANN
- SVM: only two class separation

A more separable problem: particle id with dE/dx

TPC and Vertex Detector



Variables used:

- momentum of particle
- dE/dx in Vertex Detector
- dE/dx in TPC
- standard deviations of both measurements

slightly different conditions used for SVM and ANN

- Net uses deviation of measured dE/dx from expected dE/dx curves for different particle species, whereas this a priori knowledge is not used for the SVM

- dE/dx errors not used for SVM

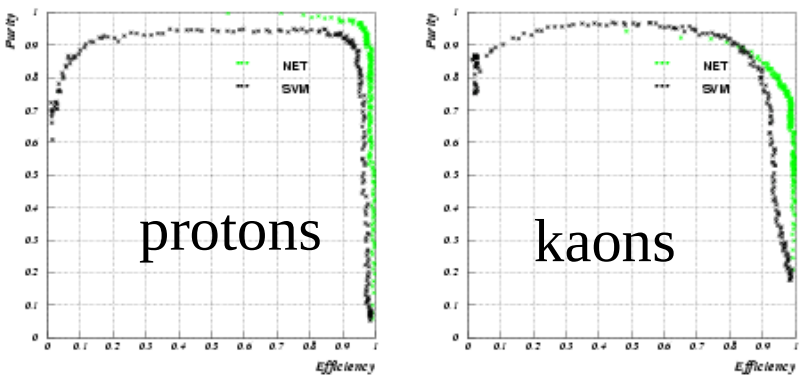
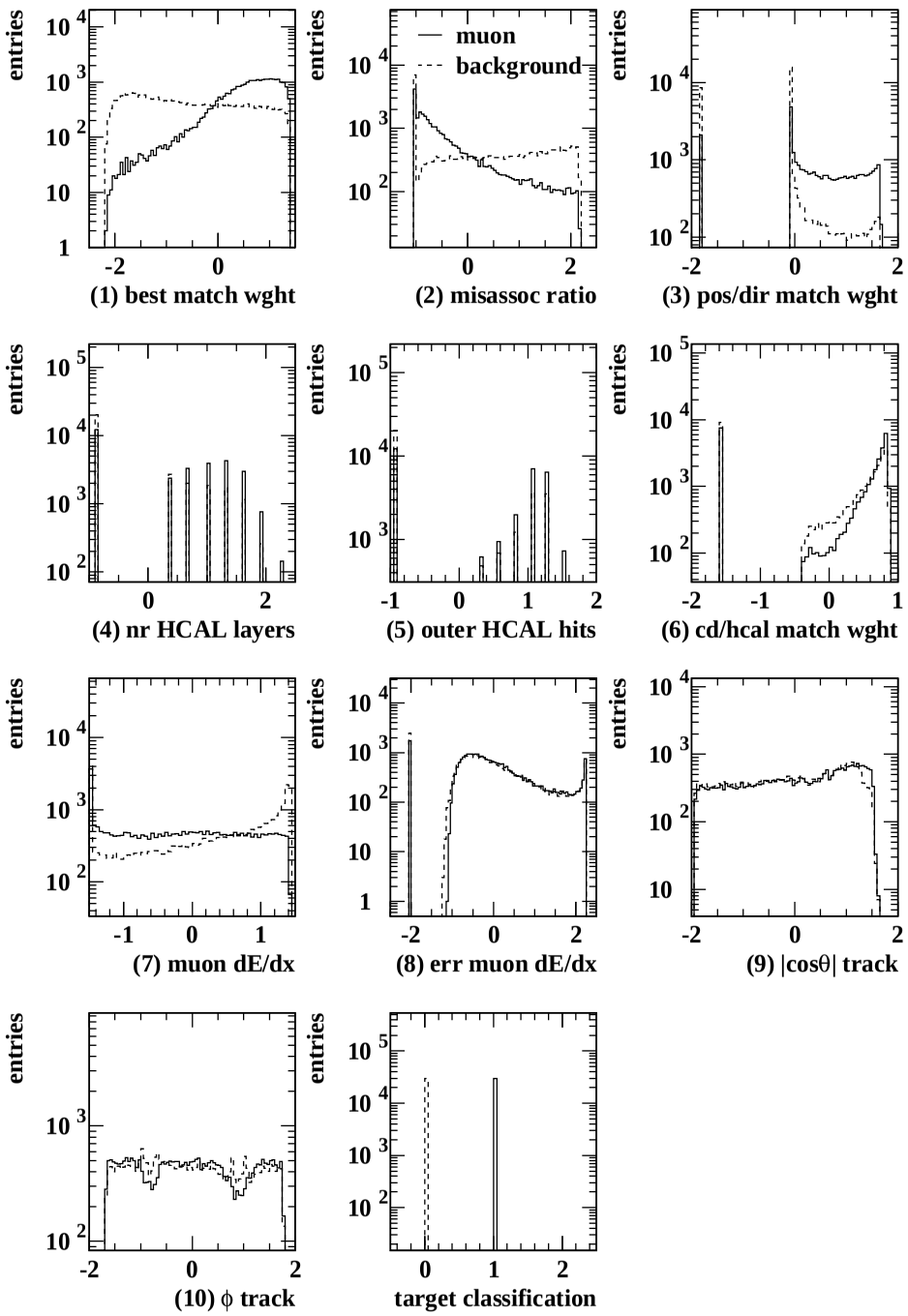


Figure 6: *Comparison between Network and SVM for protons (left) and kaons (right). For the kaon separation the SVM approaches the NN performance in some regions.*

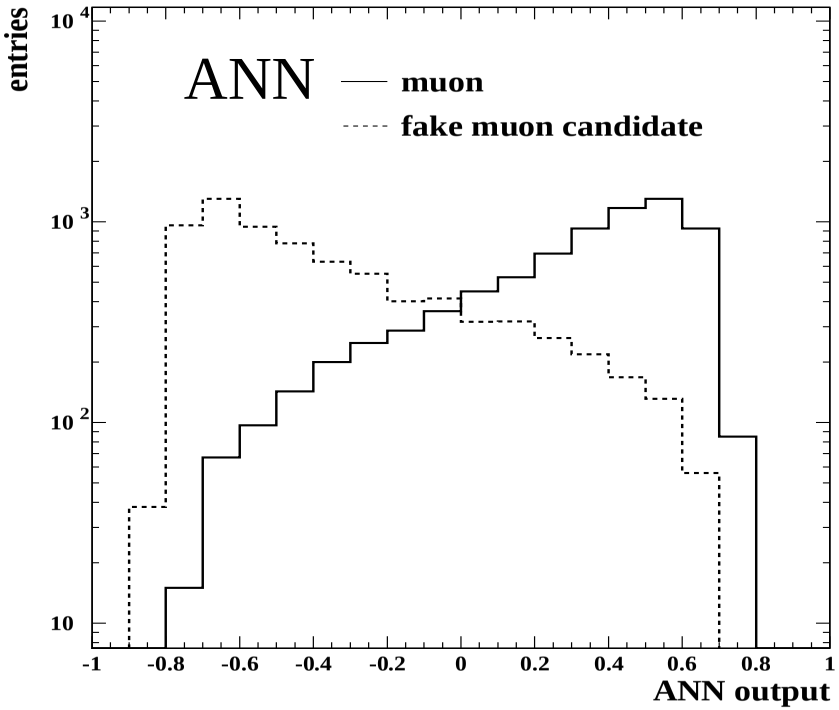
Muon identification in OPAL

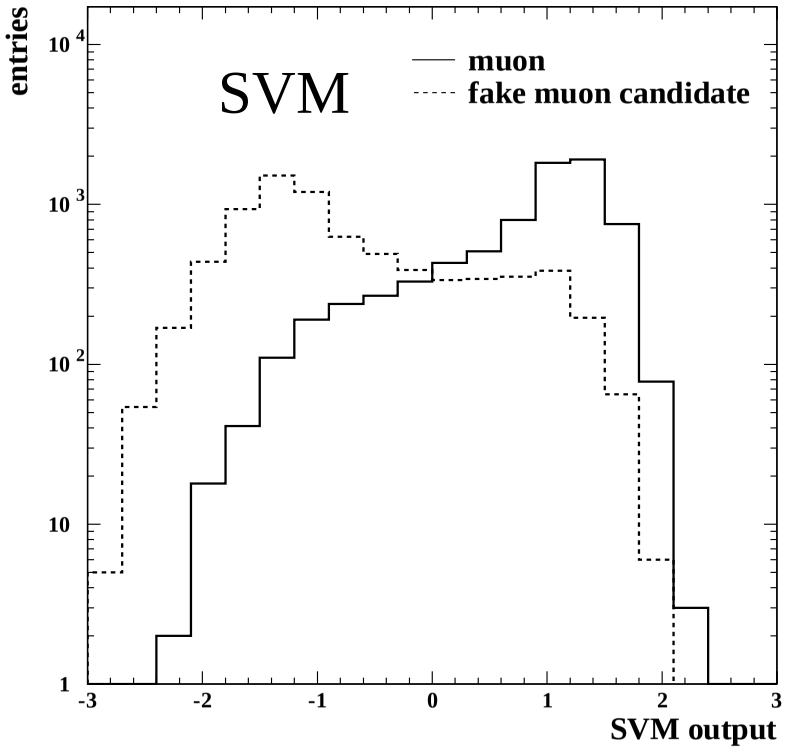


10 variables

- central track matching
- Hadronic Calorimeter (HCAL) matching
- specific energy loss (dE/dx) in jet chamber
- polar angle and phi of track

Distributions shifted and re−scaled to have mean value of 0 and standard deviation of 1





JETNET used for ANN:

3 layers

10 input nodes

18 hidden nodes

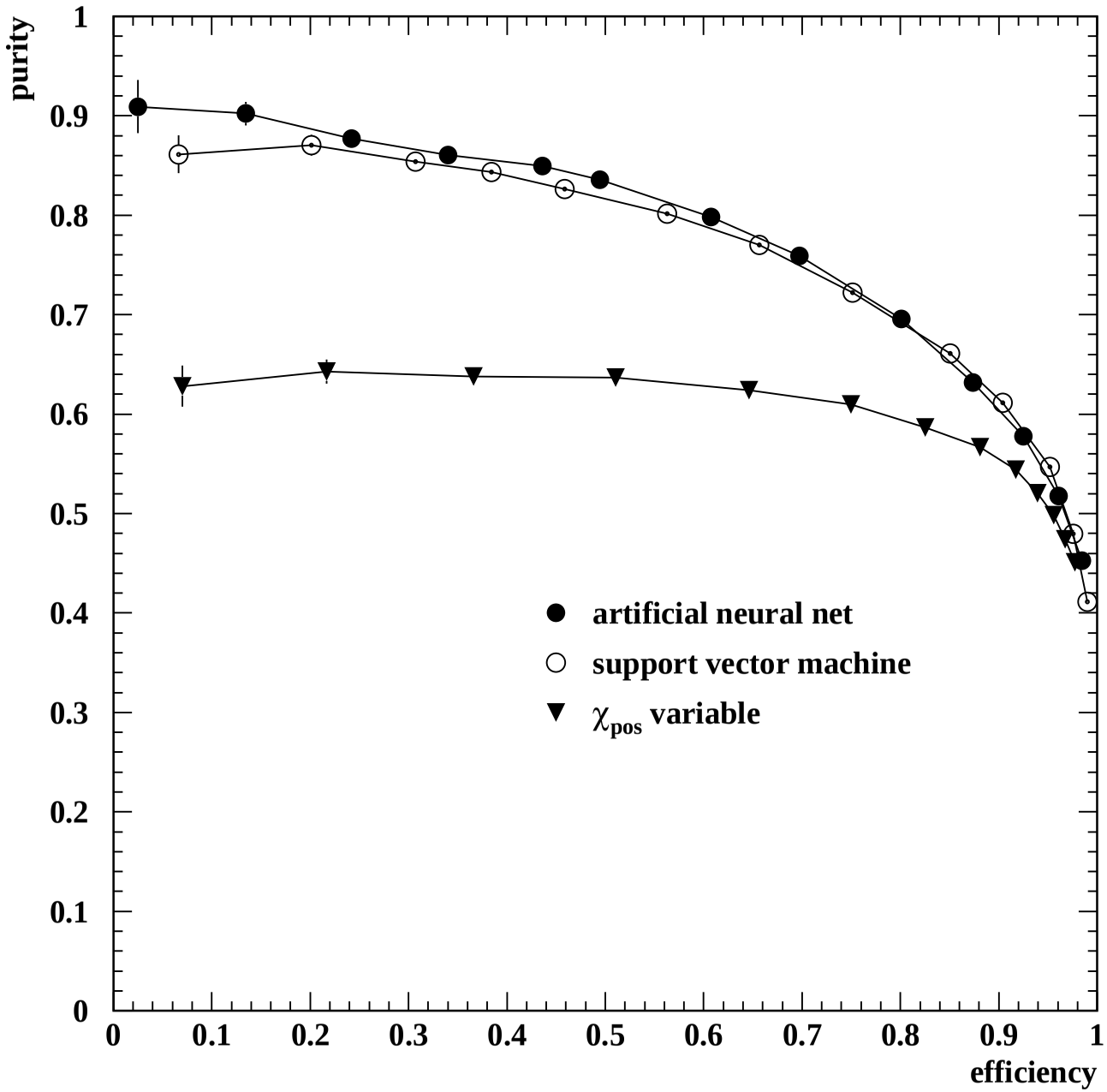
1 output node

Free parameters in SVM:

1) width of Gaussian kernel

2) regularisation parameter

Variations seems to have very little influence on SVM performance


purity
efficiency
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
artificial neural net
support vector machine
χ_{pos} variable

Charm tagging at LEP1 (hep −ex/9905027)

Identify $e^+ e^- \to c\,c$ to measure electroweak variables like R_c, A_{FB}

hemisphere tag (not event tag) using jet with highest energy

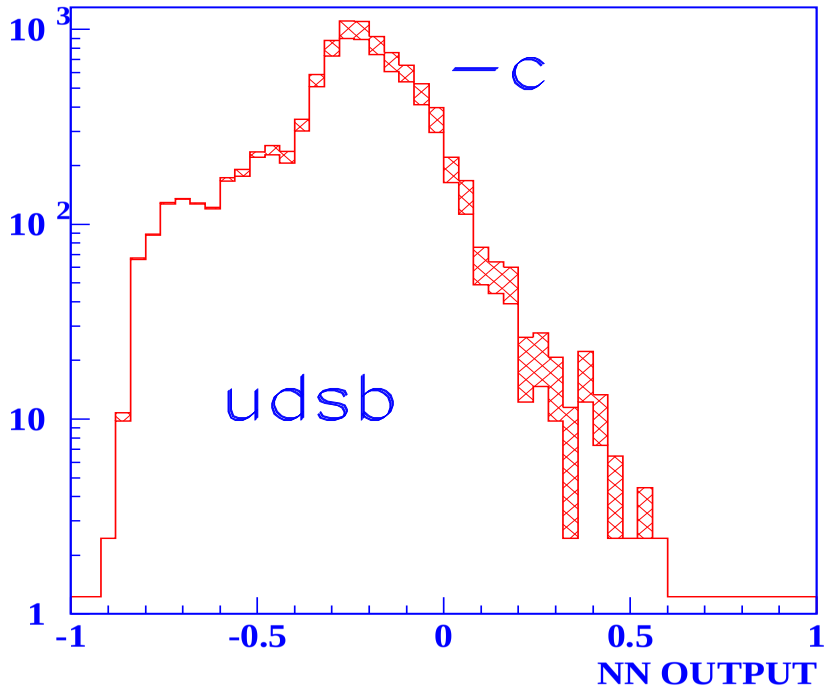
Signal lies between two background classes: uds and b

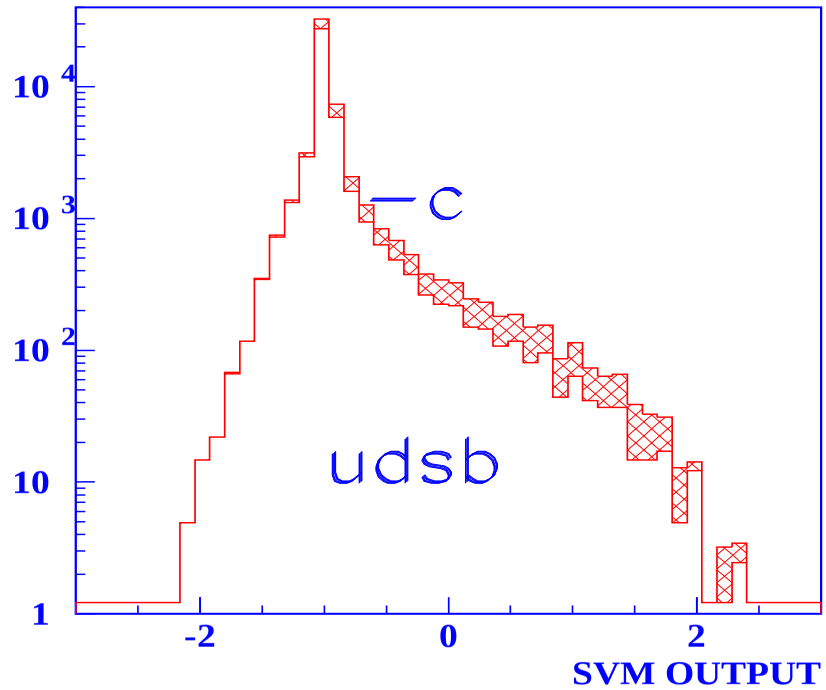
Pure charm tags (D* etc.) suffer from low statistics

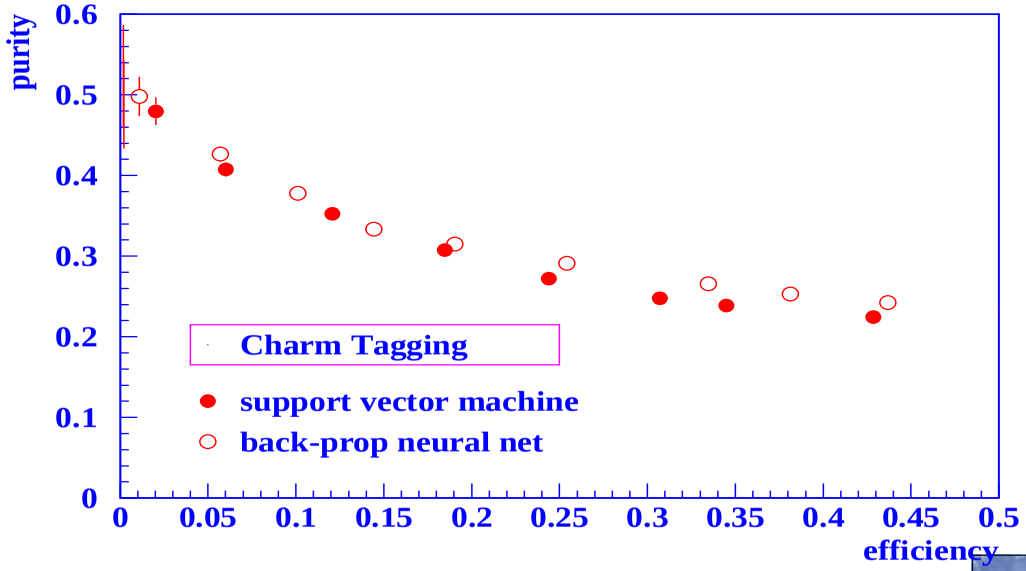
An inclusive charm tag is therefore an interesting challenge
for classifications algorithms −> common project with GMD Berlin

start with a preselected sample and 27 (!) variables:
(jet shapes, Fox−Wolfram moments, sphericity tensor...
no lepton variables to reduce correlations with lepton tags)

14 variables with largest influence in an ANN were retained







Stefan Soldner−Rembold, 3/7/2002

SUMMARY

PRO's

−no local minima when training
−few parameters
−smooth output distribution
−acceptable performace
for small training samples

Con's

−CPU time intensive
training
−slower classification
than ANN

Other issues: missing variable problem,
separation of more than two classes

Bottom line: Performance of ANN and SVM similar

Optimal selection and pre−treatment of variables very important