ALEPH 90-105
PHYSIC 90-66
J. Jousset et al.
9.7.1990

Comparison of three pattern recognition methods for the tagging of b quark events in ALEPH : Linear Discriminant Analysis, Classification Tree, Neural Network.

J. PRORIOL, J. JOUSSET, C. GUICHENEY,
A. FALVARD, P. HENRARD, D. PALLIN,
P. PERRET, F. PRULHIERE

Laboratoire de Physique Corpusculaire
de Clermont-Ferrand
IN2P3/CNRS Université Blaise Pascal
F-63177 AUBIERE CEDEX FRANCE

G. BARRAND

Laboratoire de l'accélérateur linéaire
IN2P3/CNRS Université de Paris-Sud
F-91405 ORSAY CEDEX FRANCE

ABSTRACT :

We compare the plots purity versus efficiency for the tagging of b quarks events in ALEPH with three classifiers : linear discriminant method, classification tree, neural network.

1 – Introduction

The choice of multivariate analysis methods for the pattern recognition of events in high energy physics has already been used. During the preparation of SPS [1] and LEP experiments [2] the linear discriminant analysis was used to tag top quark events. Some other studies were done [3] to recognize the number of jets of an e^+e^- event or to tag b quark events in LEP experiments. In the separation of neutral and charged clusters in ALEPH electomagnetic calorimeter the classification tree method was used too [4].

Recently [5] a comparison of linear discriminant analysis and classification tree methods have been done in ALEPH [5]. A set of Monte Carlo events reconstructed through JULIA has been used for this study.

Another successful process for event selection is the neural network method recently used [6] to tag high energy physics events.

In this paper a comparison of these three methods (Linear Discriminant Analysis, Classification Trees and Neural Network), applied to the same set of e^+e^- events is given; their efficiencies are compared.

2 – Choice of the learning sample and of the test sample

Performing a pattern recognition method needs a first step called the learning phase in which one uses some events for which the class they belong to is known. From this first step one gets the rules of classification, called classifier, an unknown event can then be classified according to such rules.

Once that the learning phase has been performed, a validation of the rules must be done via a large sample of known events, called hereafter the test sample, to minimize the statistical error.

The learning sample and the test sample are made of Monte Carlo events. These events were generated in the ALEPH Collaboration by Annecy, Clermont and Marseille. They have been reconstructed through JULIA and can thus be compared to the data. At the time no minivertex detector was implemented in the experiment.

A set of b quark events, a set of c quark events and a set of mixed u, d, s, quark events have been used, assuming for the test sample the following proportions: b : 21,9% , c : 17,1% , uds : 60,9% versus $q\bar{q}$.

In this work we consider only the case $b \rightarrow e$, the identification of the leptonic event is made with the standard subroutines of ALEPH heavy flavours group [7] to tag the $b \rightarrow e$ events.

The set of 19 genuine variables which has been used is fully described in [5].

3 – Linear discriminant analysis

The classification rule has been obtained from a learning sample of 500 b events, 500 c events and 500 uds events, giving an equal weight to the three classes.

Usually all the variables are not needed for a well doing discrimination. In this study the selection of the most discriminant variables has been done with the program SELDIS from the library MODULAD [8]; for the cases $b \to e$ five variables have been selected.

In a second step the selected variables are handled in a program of discrimination. The program DISC [8] of MODULAD allowing a discrimination between three classes has been used. We have introduced a slight modification [5] which allows to increase the purity of the classified sample of $b \to e$ events.

The test sample [5] is a set of 60 000 Monte Carlo events (20 000 for each b, c and uds quark flavours). The data sample is made of 10 000 ALEPH events; a good agreement has been found between the Monte-Carlo and the data[5].

4 – Classification tree

A second method of classification is the construction of binary trees. Such trees provide a hierarchical type of representation of the data space that can be readily used as a basis for the classification by following the appropriate branches of the tree.

We used the programs DNP (Discrimination Non Paramétrique) from MODULAD [8]. The program allows to build a large tree, then we can follow one branch of the tree and compute the purity versus the efficiency [5].

The tree can be used as a classifier. If an event belonging to an unknown class is dropped into a tree and ends up in a terminal node labelled as a given class, it is classified as an event of this class.

5 – Neural network

To get a classifier with a neural network method we have chosen for the learning process the three layers feed–forward network (input, hidden, output) with a back–propagation method commonly used [9].

To reduce the number of neurons, we have used as input variables the five variables selected by SELDIS for the linear discriminant analysis: double sphericity, momentum of the electron, transverse momentum of the electron, Fox–Wolfram coefficient H_2 and the transverse momentum of the particles of the most energetic jet of the event [5].

The network is built with 5 input neurons, 20 hidden neurons and 1 output neuron with b(event) = 1 and not b (event) = 0.

The learning sample is built with 500 b events and 500 *udsc* events. The test sample is built with 2 700 b events and 3 300 *udsc* events. We can increase the purity with the variation of the cut between 0.5 and 1 for the classification from the output neuron.

6 – Conclusion

After running the test sample events through the classifiers obtained with the three methods described here one gets a new sample of classified events. Their genuine class being known the purity versus efficiency ratio is then computed.

We have compared the three classifiers applied to the same set of Monte Carlo events. These events are in very good agreement with the ALEPH data.

The curve purity versus efficiency for the 3 classifiers is given in fig(1). The errors are statistical. One can see that the three results are comparable.

The purity of classified b samples will be improved with the new version of the ALEPH detector ; the presence of a mini vertex detector should decrease the background of light quarks.

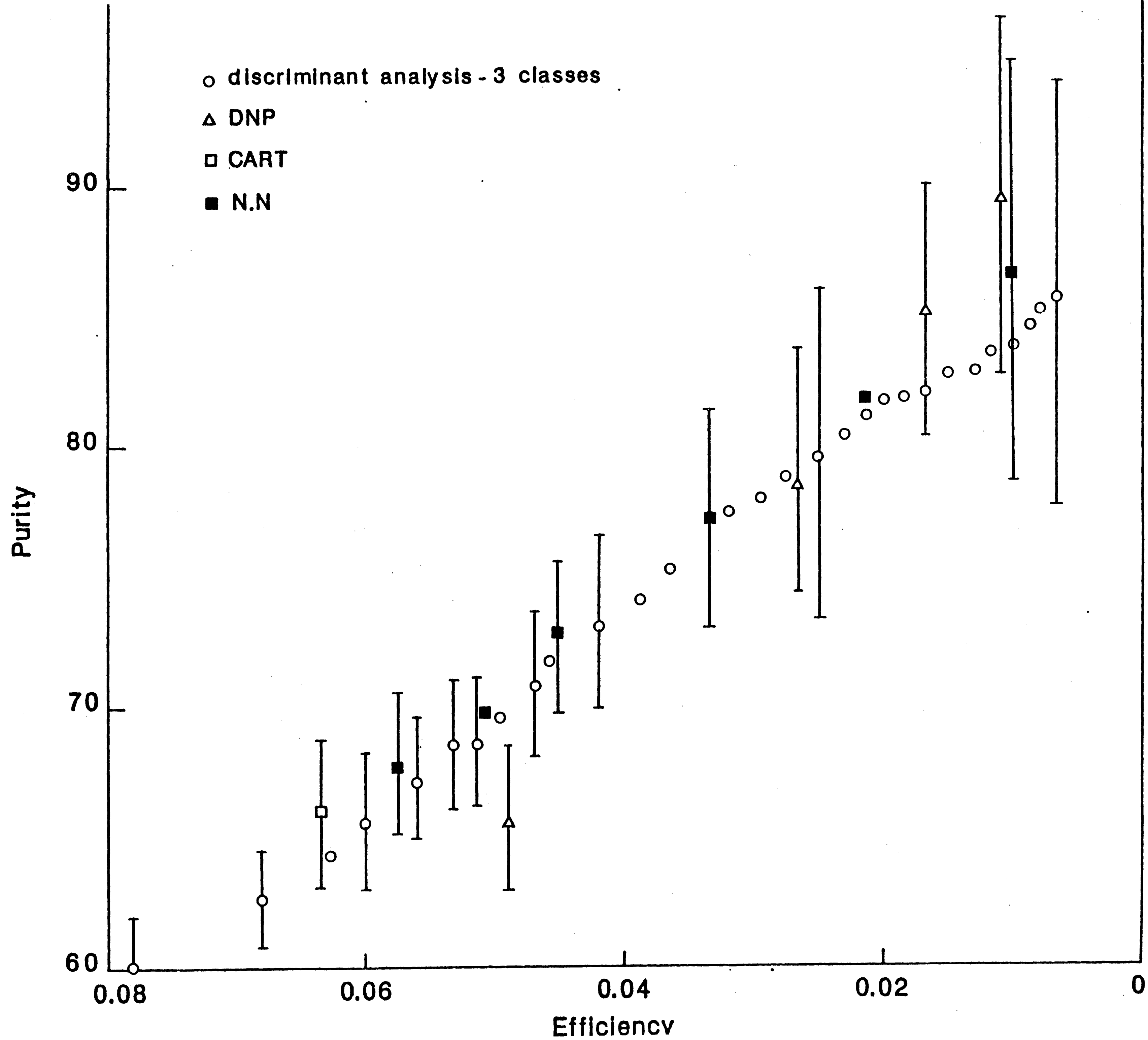
Acknowledgements

We would like to thank Drs G. CELEUX and Y. LECHEVALLIER from INRIA (Paris) for helpful comments and discussions.

Drs B. DENBY and C. PETERSON introduced some of us (J.P., J.J.) to neural network methods.

Figure caption

Purity of a classified b quark events versus efficiency. Three methods are used : linear discriminant analysis with 3 classes : 0, DNP : ⊓, neural network (5 input, neurons, 20 hidden neurons, 1 output neuron) : ⊓.



discriminant analysis - 3 classes
DNP
CART
N.N
Purity
90
80
70
60
0.08
0.06
0.04
0.02
0
Efficiency

REFERENCES

[1] R. ODORICO
Phys. Lett. B120 (1983) 219
G. BALLOCHI, R. ODORICO
Nucl. Phys. B229 (1983) 1

[2] P. LUTZ
DELPHI-NOTE 85-3 PHYS 3
D. DELIKARIS, P. LUTZ
DELPHI-NOTE 87-3 PHYS 15
D. DELIKARIS
Thèse, Collège de France (1987)
P. LUTZ
Cours de GIF (IN2P3) (1988)

[3] M. MJAED
Thèse L.P.C. Clermont (1987)
M. MJAED, J. PRORIOL
Phys. Lett. B127 (1989) 560
Ch. de la VAISSIERE Annals of the N.Y. Academic of Sciences (1989)

[4] A. EALET
Thèse Marseille (1987)
C. GUICHENEY
Rapport DEA Clermont (1989)

[5] J. PRORIOL, J. JOUSSET, C. GUICHENEY, A. FALVARD, P. HENRARD, D. PALLIN, P. PERRET, F. PRULHIERE
Aleph–Note 90-95 Physic 90-58
and International Workshop on software engineering, artificial intelligence and expert systems of high energy and Nuclear Physics, LYON–1990.
Proceedings to be published.

[6] C. PETERSON
International Workshop on software engineering, artificial intelligence and expert systems for High Energy and Nuclear Physics, LYON–1990.
Proceedings to be published.

B. DENBY
International Workshop on software engineering, artificial intelligence and expert systems for High Energy and Nuclear Physics, LYON–1990.
Proceedings to be published.

L. LÖNNBLAD, C. PETERSON, T. RÖGNVALDSSON
Preprint Lund University LU–TP 90–3

[7] ALEPH Collaboration CERN EP 90–54
P. HENRARD, A. ROUSSARIE ALEPH Collaboration.

P.PERRET, Thése Clermont 1990

[8] MODULAD INRIA Paris.

[9] D.E. RUMELHART, G.E. HINTON, R.J. WILLIAMS
Nature 323 (1986) 533

Y. Le CUN
Thèse Univ. Paris 6 (1987).