ELSEVIER

Nuclear Instruments and Methods in Physics Research A 389 (1997) 230–232

NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A

The comparison of Bayesian and neural techniques in problems of classification to multiple categories

A. Chilingarian*, S. Ter-Antonyan, A. Vardanyan

Cosmic Ray Division, Yerevan Physics Institute, Alikhanyan Brothres 2, Yerevan 36, Armenia

Abstract

The determination of elemental composition of Primary Cosmic Rays in the energy range 10^{15}–10^{17} eV is still an unsolved problem. Modern surface installation registering many characteristics of Extensive Air Shower (EAS) initiated in atmosphere by incident particles provide the possibility to handle data on an event-by-event basis and obtain results with reliability comparable with collider experiments.

We use Bayesian decision making and neural approaches for data classification into multiple categories. The Parzen window method was used for multivariate density estimation along with evolutionary algorithms for net training.

The accuracies of reconstructed proportion of different nucleus in primary flux were estimated. Both methods provide close results proving convergence to minimal achievable Bayesian risk.

1. Introduction

The ambiguity in interpreting the results of experiments with cosmic rays is connected with both significant gaps in our knowledge of the characteristics of hadron–nuclear interactions at superaccelerator energies and indefiniteness of the primary cosmic ray composition, as well as with the strong fluctuations of all the shower parameters. The extra difficulties are due to indirect experiments and hence, due to the use of Monte-Carlo simulations of development and detection of different components of nuclear electromagnetic cascades.

Only a simultaneous measurement of a large number of independent parameters in each individual event can yield reliable information to reconstruct the primary particle origin and its energetic characteristics as well as the peculiarities of strong interaction in the upper atmosphere.

To make more reliable and significant conclusions about the investigated physical phenomenon we develop a unified theory of statistical inference, based on nonparametric models, in which various nonparametric approaches (density estimation, Bayesian decision making, error rate estimation, feature extraction, sample control during handling, neural net models, etc.) are implemented.

* Corresponding author. Tel./fax: + 7 374 2 352041, e-mail: chili@yerphi.am.

The ANI program package [1] for multidimensional data analysis and non-parametric statistical inference was used for estimating possible accuracies of primary reconstruction using the simulation data obtain by CORSIKA [2] code.

2. Bayesian and neural decision rules

The Bayesian approach in classifying a mixture of distribution consists in the minimization of the losses due to incorrect classification [3]. Therefore, the Bayesian decision rule takes the form (a simple loss function is assumed)

$$A = \operatorname{argmax}_i \{\hat{p}(A_i/v)\}, \quad i = 1, \dots, L, \tag{1}$$

where the space of possible states of nature $A \equiv (\mathrm{p}, \alpha, \mathrm{O}, \mathrm{N}, \mathrm{Fe})$ consists of 5 groups of primary nucleous, $\boldsymbol{v}$ is a multivariate measurement and $\hat{p}(A_i/v)$ are nonparametric estimates of a posterior densities, connected with conditional ones (obtained from simulation) by Bayes theorem:

$$\hat{p}(A_i/v) = \frac{P_i\,\hat{p}(v_i/A_i)}{p(v)}. \tag{2}$$

Conditional densities are estimated by training samples using Parzen's method with automatic kernel width adaptation. In this method some probability density values are calculated which correspond to different values of

PII S0168-9002(97)00134-4

method parameters. Then the sequence obtained is ordered and the median of this sequence is chosen as final estimate. The probability density is estimated by

$$\hat{p}(v/A_i) = \frac{W_j(\det R)^{-1/2}}{2\pi^{d/2}h^d} \sum_j^{M_i} e^{-r_j^2/h^2}, \quad i = 1, \ldots, L, \tag{3}$$

where d is the feature space dimensionality, M is the number of events in the ith TS class, r_j is the distance to the jth neighbour in the Mahalanobismetric:

$$r_j = (v - u_j)^T R^{-1} (v - u_j), \tag{4}$$

where R is a sampling covariance matrix of the class to which $\boldsymbol{u}$ belongs. W_j are the event weight, h is width of the kernel.

The Neural decision making is another nonparametric technique, mapping the multidimensional measurements into one-dimensional "decision" (0–1) interval [4]. The "target" output $\mathrm{OUT}^{\mathrm{target}}(k)$ for events of the kth category events (we restrict ourselves to networks with a single output node) is determined to maximize the shift of the alternative classes from each other:

$$\mathrm{OUT}^{\mathrm{target}}(k) = \frac{k-1}{K-1}, \quad k = 1, K, \tag{5}$$

where K is the total number of classes, 5 in our case. The actual events classification is performed by comparing the obtained output value with the predefined intervals in the (0, 1) interval. We expect that the data flow passing through the trained net will be divided in 5 clusters concentrated in the different regions of the (0, 1) interval (see Fig. 1).

This neural decision rule is also a Bayesian one, therefore the output signal of a properly trained feedforward neural net is an estimate of the a posteriori probability density [5].

The expected minimal classification errors caused by the overlap of the distributions (the Bayes error) depends on the discriminative power of the feature subset selected and on the learning power. By moving the decision points along the (0, 1) interval we can change the relation between the errors of the first and second kind (the position of the decision points is the neural analog of the loss function in the Bayesian approach).

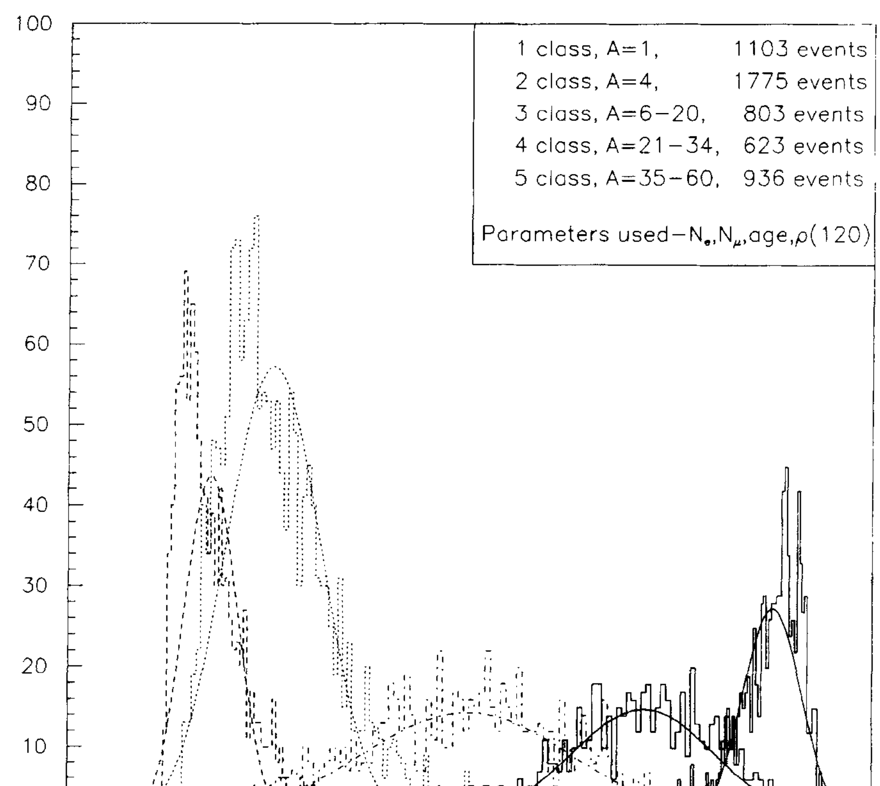

Fig. 1. N.N. classification of EAS registered at 3250 m a.s.l. (CORSIKA M.C.) in the energy range 10^{14}–10^{16} eV.

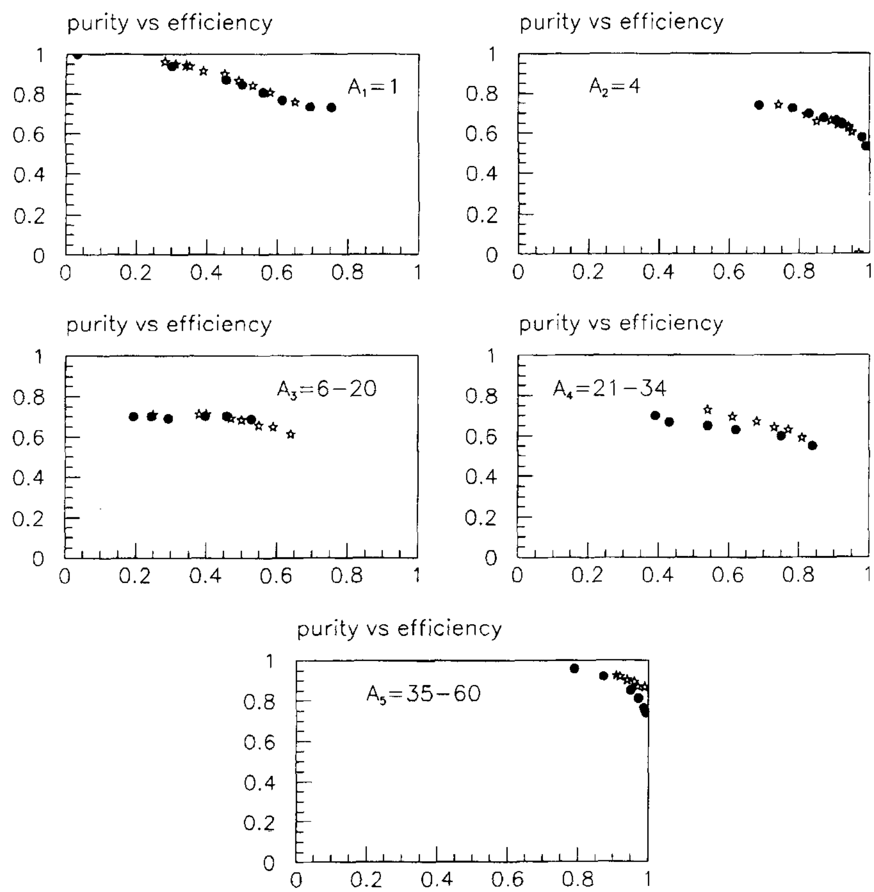

Fig. 2. N.N. classification of EAS registered at 3250 m a.s.l. (CORSIKA M.C) in the energy range 10^{14}–10^{16} eV. (●) Neural Net. (☆) Bayes.

3. Results

The primary nucleus estimation was done for ANI experiment now operating on Aragatz mountain research station of Yerevan Physics Institute, Armenia. The experimental complex consists of a ground-based shower array, EAS muon detectors installed in the underground hall and a precise electron density detector. The location of station (3200 m above sea level) permits the accurate estimation of the energy of a primary particle [6], also providing opportunity to obtain rather good accuracies of primary-type estimation.

The CORSIKA simulation were used for training of Bayesian and Neural algorithms. From the purity-efficiency plots (Fig. 2), one can see the good agreement of both nonparametric approaches. Different points were obtained by altering the a priori probabilities in Bayesian method and decision points–in neural method.

Acknowledgements

We wish to acknowledge the members of the KASCADE collaborations for the use of the CORSIKA EAS simulation program.

We thank Prof. Dr. H. Rebel, Dr. J. Knapp and M. Roth for valuable remarks concerning new methods of data analysis. This work was partly supported by the research grant No. 94694 by the Armenian government.

References

[1] A.A. Chilingaryan, Comput. Phys. Commun. 54 (1989) 381.
[2] Capdevielle et al., Kernforschungszentrum, Karlsruhe, report 4998, 1992.
[3] A.A. Chilingarian, Pattern Recognition Lett. 16 (1995) 333.
[4] A.A. Chilingarian, Neurocomputing 6 (1994) 497.
[5] D.W. Ruck, K.S. Rogers et al., IEEE Trans. Neural Networks 1 (1990) 296.
[6] R.M. Martirosov, J. Procureur and J.N. Stamenov, Nuovo Cimento 108 (1995) 299.