The Stellar Content of Local Group Galaxies
IAU Symposium, Vol. 192, 1999
Patricia Whitelock and Russell Cannon, eds.

The Application of Artificial Neural Networks: A Catalog of Spectral Indices

R.K. Gulati

INAOE, Apartado postal 51 y 216, 72000 Puebla, Pue, Mexico

A. Bravo and G. Padilla

Instituto Technológico de Apizaco, Apizaco, Tlax., Mexico

L. Altamirano R.

INAOE, Apartado postal 51 y 216, 72000 Puebla, Pue, Mexico

Abstract. Artificial neural networks have been applied to a catalog of spectral indices based on low-resolution spectra in the wavelength regions 3820–4500 Å and 4780–5450 Å. We classify the spectral indices using supervised neural networks. This project is in continuation of the previous efforts by one of us, applying new tools to determine basic properties of stars. We envisage the use of such methods for stellar contents studies of local stellar systems for which spectroscopic surveys are underway.

1. Introduction

With the improvements in observational tools, it is feasible to have larger databases of low-resolution spectra. To extract observational and physical properties from such larger databases, it is necessary to employ computer-based methods, namely artificial neural networks (ANNs) and principal component analysis. Various groups have applied these tools to classify galaxy spectra (Folkes et al. 1996) and stellar spectra (von Hippel et al. 1994; Weaver & Torres-Dodgen 1997; Bailer-Jones et al. 1998). One of us has already applied these tools to classify ultraviolet (UV) and optical spectra (Gulati et al. 1994a; Gulati et al. 1994b; Singh et al. 1998); to extract reddening properties from UV spectra (Gulati et al. 1997a); and to compare observed spectra with theoretical ones (Gulati et al. 1997b). We present here another application of ANNs to classify a larger database of spectral indices based on low-resolution spectra. In section 2 we briefly discuss underlying principal of the method; and in section 3 we present the application of the method to a catalog of spectral indices.

2. Method

In astronomy it is common practice to fit a model with several free parameters in order to extract useful information from observations. This is often done by means of χ^2 minimisation. A new way to derive model parameters is by

https://doi.org/10.1017/S0074180900204373 Published online by Cambridge University Press

using supervised artificial neural networks. In this case input information is mapped into the output parameters through a non-linear function with weights as free parameters. These free parameters are determined by minimising the cost function:

$$E = \frac{1}{2} \sum [T_i - f(\mathbf{w}, \mathbf{x_i})]^2$$

where the T_i are target parameters and the $\mathbf{x_i}$ are input information for the ith pattern. Contrary to looking at the variance between a set of test data and a template data set, the variance between the target and network parameters is minimised here. The minimisation is done during "training" of the network by using a back propagation algorithm (see also Rumelhart et al. 1986). Once the network has converged to a given threshold value, the weights of the converged network are used to determine parameters for a test data in terms of a trained data set.

3. Application

3.1. Spectral Classification Based on Spectral Indices

Jones (1996) has provided a catalog of spectral indices based on the set of spectra which he observed at the KPNO observatory with the coudé feed telescope. The indices were measured, according to the definitions of Rose (1994) and Worthey et al. (1994), from the spectral regions 3820–4500 Å and 4780–5450 Å; they involve 32 indices in the former and 16 indices in the latter wavelength regions. We used the indices, together with the spectral classes given in his catalog, to apply ANNs for spectral classification. Details about the input data and method will be described elsewhere (Gulati et al. 1998).

While designing the architectures for supervised neural networks, one needs a minimum of three layers. The first layer is used to feed input data, the second is required to process the data and the third gives the output classes. In this application, we designed two networks, one with 32 input nodes from the regions around 4000 Å and the other with 16 nodes from the regions around 5000 Å. The number of hidden nodes in the hidden layer was set to optimal size so that the networks can generalize classification of the test patterns. In the following section we present the results based on the above networks.

3.2. Results Applying ANNs

In Fig. 1(a), we show a correlation plot for the set of training patterns selected from the catalog each with 16 spectral indices (case I). Fig. 1(c) is the same but for the 32 spectral indices (case II), while in Figs. 1(b) and 1(d) we show the performance of the networks by plotting network class versus catalog class for the test patterns, with 16 and 32 indices respectively. Lines at 45° are drawn to show the ideal correlation between the two classification schemes. Non-weighted linear least squares fits to the data in Figs. 1(a)–(d) suggest a tight correlation between the catalog class and the network class for both cases. However, the network trained with the 32 indices performs better on the test samples than that with 16 indices, as indicated in the global error (RMSE) of the fitting

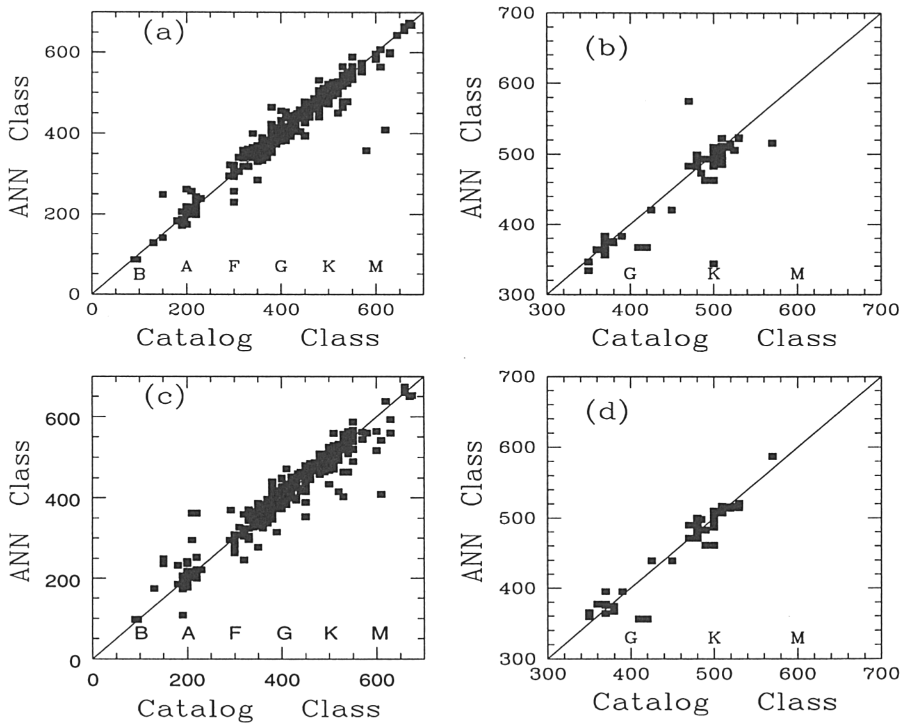

Figure 1. Correlation of the network classifications to the catalog classifications for a training set each with 16 spectral indices (a). The codes used to represent the classes are shown along the absicassae. Same for the case 32 spectral indices is shown in (c). Comparison of the network classifications with the catalog classifications for the test set with 16 and 32 indices is shown in (b) and (d) respectively.

straight line (2 subclasses in case II vs. 3 subclasses in case I). Note that the fitted slope in each figure is close to the ideal correlation. There are a few spurious points which lie away from the the ideal correlation and the causes of their discrepancies need to be investigated further.

We illustrate in this paper that a catalog of spectral indices built for population synthesis work can be utilized to determine basic properties of stars. We envisage the use of artificial neural networks for deriving observational properties from large spectroscopic surveys of stellar spectra carried out with multi-object spectrographs.

Acknowledgements. R.K. Gulati acknowledges the grant from CONACyT for the support of this project.

References

Bailer-Jones, C.A.L., Irwin, M., Von Hippel, T. 1998, MNRAS, 298, 361

Folkes, S.R., Lahav, O., Maddox, S.J. 1996, MNRAS, 283, 651

Gulati, R.K., Gupta, R., Gothoskar, P., Khobragade, S. 1994a, ApJ, 426, 340

Gulati, R.K., Gupta, R., Gothoskar, P., Khobragade, S. 1994b, Vistas in Astro., 38(3), 293

Gulati, R.K., Gupta, R., Rao, N.K. 1997a, A&A, 322, 933 (G97a)

Gulati, R.K., Gupta, R., Singh, H.P. 1997b, PASP, 110, 843 (G97b)

Gulati, R.K., Altamirano, R., Bravo, A., Padillo, G. 1998, in preparation

von Hippel, T., Storrie-Lombardi, L.J., Storrie-Lombardi, M.C., Irwin, M.J. 1994, MNRAS, 269, 97

Jones, L.A. 1996, Ph.D. Thesis, North Carolina University, Chapel Hill

Rumelhart, D.E., Hinton, G.E., Williams, R.J. 1986, Nature, 323, 533

Rose, J.A. 1994, AJ, 107, 206

Singh, H.P., Gulati, R.K., Gupta, R. 1998, MNRAS, 295, 312

Weaver, Wm.B., Torres-Dodgen, A.V. 1997, ApJ, 487, 847

Worthey, G., Faber, S.M., Gonzalez, J.J., Burstein, D. 1994, ApJS, 94, 687