Astronomical Data Analysis Software and Systems VII
ASP Conference Series, Vol. 145, 1998
R. Albrecht, R. N. Hook and H. A. Bushouse, eds.

Automated Spectral Classification Using Neural Networks

E. F. Vieira
Laboratorio de Astrofísica Espacial y Física Fundamental
P.O.Box 50727, 28080 Madrid, Spain, Email: efv@laeff.esa.es

J. D. Ponz
GSED/ESA, Villafranca
P.O.Box 50727, 28080 Madrid, Spain, Email: jdp@vilspa.esa.es

Abstract. We have explored two automated classification methods: supervised classification using Artificial Neural Networks (ANN) and unsupervised classification using Self Organized Maps (SOM). These methods are used to classify IUE low-dispersion spectra of normal stars with spectral types ranging from O3 to G5.

1. Introduction

This paper describes the application of automated methods to the problem of the classification of stellar spectra. The availability of the IUE low-dispersion archive (Wamsteker et al. 1989) allows the application of pattern recognition methods to explore the ultraviolet domain. The analysis of this archive is especially interesting, due to the homogeneity of the sample.

The present work has been done within the context of the IUE Final Archive project, to provide an efficient and objective classification procedure to analyze the complete IUE database, based on methods that do not require prior knowledge about the object to be classified. Two methods are compared: a supervised ANN classifier and an unsupervised Self Organized Map (SOM) classifier.

2. The Data Set

The spectra were taken from the IUE Low-Dispersion Reference Atlas of Normal Stars (Heck et al. 1983), covering the wavelength range from 1150 to 3200 Å. The Atlas contains 229 normal stars distributed from the spectral type O3 to K0, that were classified manually, following a classical morphological approach (Jaschek & Jaschek 1984), based on UV criteria alone.

The actual input set was obtained by merging together data from the two IUE cameras, sampled at a uniform wavelength step of 2 Å, after processing with the standard calibration pipeline. Although the spectra are good in quality, there are two aspects that seriously hinder the automated classification: interstellar extinction and contamination with geo-coronal Ly-α emission. Some pre-processing was required to eliminate these effects and to normalize the data.

All spectra were corrected for interstellar extinction by using Seaton's (1979) extinction law. Figure 1 shows original and corrected spectra, corresponding to a O4 star; the wavelength range used in the classification is indicated by the solid line.

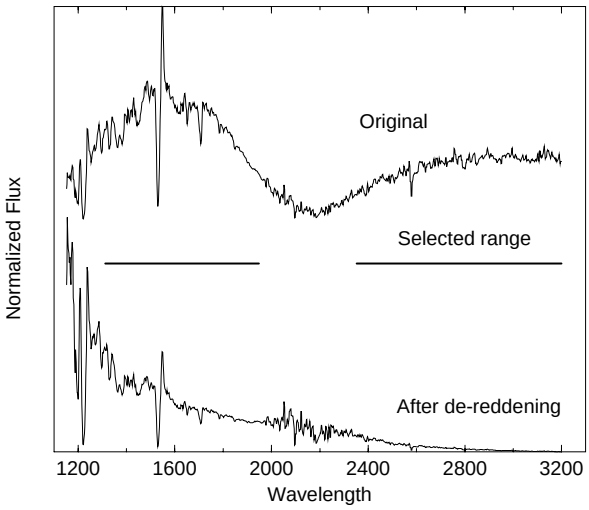

Figure 1. Original and de-reddened spectra.

3. Supervised Classification Using ANN

A supervised classification scheme based on artificial neural networks (ANN) has been used. This technique was originally developed by McCullogh and Pitts (1943) and has been generalized with an algorithm for training networks having multiple layers, known as back-propagation (Rumelhart et al. 1986).

The complete sample in the Atlas was divided into two sets: 64 standard stars, with spectral types from O3 to G5, was used as the training set. The remaining spectra were used as a test to exercise the classification algorithm. The network contains $744 \times 120 \times 120 \times 51$ neurons. The resulting classification error on the test set was 1.1 spectral subclasses. Figure 2 shows the classification diagrams, comparing automatic classification (ANN) with manual (Atlas) and with a simple metric distance algorithm.

4. Unsupervised Classification Using SOM

In the Self Organized Map (SOM) the net organizes the spectra into clusters based on similarities using a metric to define the distance between two spectra. The algorithm used to perform such clustering was developed by Kohonen (1984).

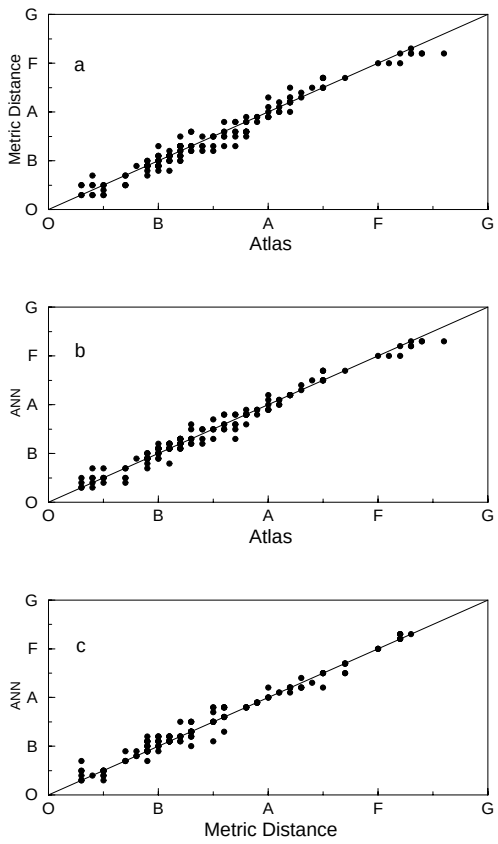

Figure 2. Results of supervised classification.

A 8×8 map with 744 neurons in the input layer was exercised on the same input sample. The training set was used to define the spectral types associated to the elements in the map. This classifier gives an error of 1.62 subclasses when compared with the Atlas, with a correlation of 0.9844. In addition, 27 stars could not be classified according to the classification criterion used in this experiment. Figure 3 shows the classification diagrams, comparing the SOM classifier with ANN and manual classification.

5. Conclusions

Two automated classification algorithms were applied to a well defined sample of spectra with very good results. The error found for supervised algorithm is 1.10 subclasses and 1.62 subclasses for the unsupervised method.

These methods can be directly applied to the set of spectra, without previous analysis of spectral features.

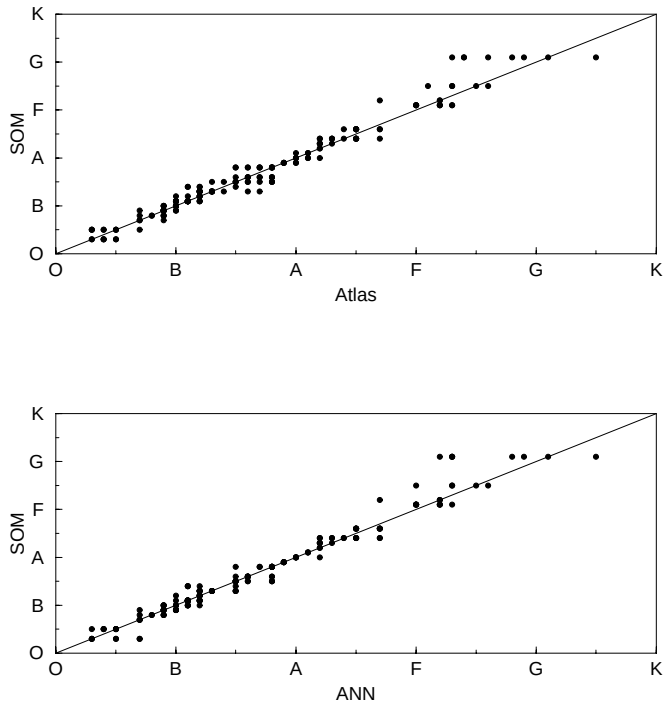

Figure 3. Results of unsupervised classification.

References

Heck, A., Egret, D., Jaschek, M., & Jaschek, C. 1983, (Feb), IUE Low Dispersion Spectra Reference Atlas, SP 1052, ESA, Part1. Normal Stars

Jaschek, M., & Jaschek, C. 1984, in The MK process and Stellar Classification, David Dunlap Observatory, 290

Kohonen, T. 1984, Self-organization and Associative Memory Volume 8 of Springer Series in Information Sciences (Springer Verlag, Nueva York)

McCullogh, W.S., & Pitts, W.H. 1943, Bull. Math. Biophysics, 5, 115

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

Seaton, M.J. 1979, MNRAS, 187, 73P

Vieira, E.F., Ponz, J.D. 1995, A&AS, 111, 393

Wamsteker, W., Driessen, C., & Muñoz, J.R. et al. 1989, A&AS, 79, 1