Bull. Astr. Soc. India (1996) **24**, 21–30

Classification of ultraviolet stellar spectra using artificial neural networks

R.K. Gulati and Ranjan Gupta

IUCAA, Post Bag 4, Ganeshkhind, Pune 411007, India

Pradeep Gothoskar and Shyam Khobragade

NCRA, TIFR Centre, P.O.Box 3, Pune 411007, India

Received 24 January 1995; accepted 29 May 1995

Abstract. Scheme based on the artificial neural networks (ANNs) trained with multi-layer back propagation using single level network for classifying ultraviolet spectra in the *IUE Low-Dispersion Spectra Reference Atlas* is presented. We compare the classification results obtained from single-level network with the one from multi-level neural tree network. Both the schemes could successfully classify the test set of spectra in terms of spectral type as well as luminosity.

Key words : stars: stellar classification; galaxies: stellar content; methods : statistical

1. Introduction

Progress in ground based and space instrumentation has brought us to a new era of spectroscopy. Large stellar spectral databases through data archives have started becoming available in the wide wavelength range. One of the main tasks is to develop machine-based methods to analyze these spectra and to infer the physical and chemical properties of the stars for understanding the evolutionary processes in the solar neighbourhood and other stellar systems. In principle, these methods should have features like objectivity, quantifiability, and repeatability, which need to be tested on the data set with known parameters before applying to large databases. The conventional way is to classify the stellar spectra in the light of common visible properties, and a comprehensive review on this subject has been made by Jaschek & Jaschek (1990). In fact the spectral classification is a useful tool, not only to arrange the spectra of the stars into meaningful and self consistent groups, but also for providing information about the astrophysical parameters, namely effective temperature, luminosity, mass and radius etc. The distribution of the stars in the various groups can be used for the statistical study of stellar populations in the solar neighbourhood and for other stellar systems which are increasingly becoming accessible through spectroscopic studies.

Conventionally, stellar classification used to be done by the visual inspection of the individual stellar spectra by the human experts and this approach led to many non-unique data bases with inherent biases of the human classifiers. In contrast to this manual scheme of classification, the automated classification schemes have numerous advantages viz. high speed, on line classification, homogeneity (elimination of personal error), accuracy, detection of variability and possibility of classification of higher dimension.

There are two approaches to perform the automated classification: criteria evaluation and pattern recognition. In the first case, a set of specified criteria, such as the width or strength of specified lines, are automatically measured and then calibrated in terms of the desired quantities (e.g., spectral types and luminosity class).

On the other hand, in pattern recognition an observed spectrum is compared with a library of reference spectra following a cross-correlation scheme or multi variant statistical methods. These techniques have their limitations, for example the cross-correlation technique (Adorf 1986) is limited due to the fact that the continuum dominates the classifier. Moreover, most of these techniques are based on linear operations whereas the inter-relation among the parameters affecting the stellar spectra can actually be non-linear and thus one requires a technique which can handle this type of relationship. The Artificial Neural Network (ANN) schemes, which we are using to perform spectral classification, are nonlinear extensions of statistical methods and the reader is referred to the special issue on "Neural Networks in Astronomy" (Lahav & Lombardi 1994).

The schemes discussed above have to be automated for projects involving classification of large databases as these require higher degree of efficiency which has to be achieved without any manual intervention. The Anglo-Australian Telescope (AAT) multi-object spectrograph will provide lots of new digital data and perhaps a "quick-look" approach could be provided automatically followed by a more careful study of the most interesting stars. Automated stellar spectral classification using ANN technique has been recently achieved in the optical and the near infra-red (NIR) wavelength region in recent times (von Hippel *et al.*, (1994); Gulati *et al.*, (1994a), hereafter referred to as G94a, and Weaver 1994). However in the ultra-violet (UV) region, till recently, only conventional statistical methods of classification have been tested on existing catalogs (Jaschek & Jaschek 1984; and Heck *et al.* 1986). In a preliminary study we (Gulati *et al.*, 1994, hereafter referred to as G94b)[1] have attempted an Artificial Neural Network model based on Multilevel Tree Network (MTN) and demonstrated its performance as compared to catalog classification. In this paper we compare MTNs performance with the model similar to the one used in the optical region (G94a) for a limited database.

The UV region of the spectrum is very important in the understanding of the early type stars as they exhibit most of these fluxes in the UV. Further the comparison of the UV spectral

[1] Detailed description of MTN is given in Vistas in Astronomy paper.

classification with that of its optical counterpart can resolve the anomalies of classification and supplement information on interstellar extinction (Gulati *et al.* 1987), Moreover, detectors have been vastly improved, which in turn allow access to wavelength regions which were not possible to observe earlier. In the ultra violet (UV) region the spectral data has been gathered with satellite based platforms, of which the IUE satellite has been the most successful mission. For details on the IUE instruments readers are referred to the article by (Boggess *et al.* 1978a). This satellite has been providing us with homogeneous library of UV stellar spectra in low and high dispersions. It is capable of observing distant stellar objects which would otherwise require much larger ground based telescope times.

The paper is organized as follows: In section 2 we discuss the input data and its pre-processing, while in section 3 the classification scheme of ANNs is presented. Section 4 gives the results and discussion, and finally section 5 contains the conclusions.

2. The input data and pre-processing

The database used for this work is the 'IUE Low resolution spectra reference atlas, Normal Stars; ESA SP-1052' by (Heck *et al.* 1984) which has 229 low dispersion spectra (spanning O to K spectral types) collected by the International Ultraviolet Explorer (IUE) satellite. Since this catalog contains biased samples of different spectral classes with insufficient late type spectra, we have restricted ourselves from O to early F type spectra. Further the luminosity classes have been restricted to super-giants, giants and dwarfs due to inadequate samples for finer subdivisions of luminosity classes. This criterion has led us to select a total of 211 spectra from this catalog, of which a subset of 128 spectra spanning 75 spectro-luminosity classes (O to F) form the training set and the remaining 83 form the test set. The selection of the training set was made by avoiding as far as possible, the cases where interstellar extinction was high so that the interstellar reddening does not affect the continuum. For the proper training of the neural network, each spectro-luminosity class must have at least two patterns (i.e. 150 patterns for training) and for this purpose some spectro-luminosity types which had only one example had to be duplicated for the training session. The catalog classification for the spectra were taken from (Heck *et al.* 1984) where the UV spectral classification is given in terms of O, B, A and F (main classes), sub-classes (0.0 to 9.5) and luminosity (super-giants, giant, dwarfs etc.). The spectra cover the range of 1213-3201 Å with one sample per 2 Å (spectral resolution of ~ 6 Å) i.e. 995 data points. All the spectra were normalized such that the maximum flux in each corresponds to unity for the sake of uniformity and this ensures that the line and continuum information are included in the sample.

Conventionally, spectro-luminosity class is named in an alpha-numeric fashion and by using an integer code one can simplify the process of data-coding into a computer. We have coded the spectro-luminosity classes into a four-digit number for the purpose of quantitative analysis of the classification as follows :

$$\text{CODE NUMBER} = 1000 \times A1 + 100 \times A2 + A3 \tag{1}$$

where A1 is the main spectral type of the star (i.e. O to F types coded as 1 to 4), A2 is the sub-spectral type of the star (coded from 0.0 to 9.5) and A3 the luminosity class of the star (i.e. 2, 5 or 8 for s, g or d). As an example, in our scheme of numbering, stars having spectro-luminosity classes dB2.5, gO9.5 and sF7 (ultraviolet classes) would be coded as 2258, 1955 and 4702 respectively.

Each spectro-luminosity class is characterized by a set of absorption features appearing at a few known wavelengths which are also the diagnostic of the spectral type (Heck *et al.*, (1986), Table 1), and therefore useful for the UV spectral classification. We have used the line depths at these selected wavelengths (a total of 35 different positions within the spectral range of interest). This improves the efficiency of the network as compared to allowing the full spectra of 995 data points for training the ANN. The typical stellar UV spectra for O, B, A and F type stars along with the wavelength positions where the line depths (flux values) are monitored for ANN training and test sessions is shown in Fig. 1 of G94b. In this figure we have purposely offset each spectrum in order to display them without any overlapping and thus the intensity scale is set to arbitrary units. It is evident from the figure that the early type stars have higher flux values in the shorter wavelength region but the number of spectral features increases and tends to concentrate in the longer wavelength region for the late type stars.

3. Artificial Neural Networks scheme (ANNs)

Artificial neural networks (hereafter, ANNs) are computational models derived from the simulation of the human brain (McCulloch & Pitts 1943; Hopfield & Tank 1986). These networks consist of a series of processing elements, known as *neurons* or *nodes*, arranged in layers. These nodes are analogous to biological neurons with inputs (dendrites), summing nodes (soma or cell bodies), outputs (axons) and weighted connections between layers (synapses).

These networks have ability to learn from examples and store the information in weights in a generalized form. This facilitates the appropriate classification of the input patterns which are not included in the training set, provided the training set covered the representative patterns for all the classes. The features of ANN, learning and generalization, thus enable these networks for solving imaging and signal processing tasks which cannot be readily attacked using rule-based or other conventional techniques.

Various ANN architectures and learning algorithms have been developed and are being used in academic research as well as in industrial applications. Plenty of books and research papers on this topic are available. A few network architectures and learning rules have been reviewed by (Adorf 1989; Widrow & Lehr 1990; Clark 1991).

Recently, ANN algorithms are being used for various astronomy applications, star-galaxy separation (Odewahn *et al.*, 1992), stellar spectral classification (von Hippel *et al.*, (1994), G94a, and G94b, Weaver & Torres-Dodgen 1995; Vieira & Ponz 1995), and morphological classification of galaxies (Storrie-Lombardi *et al.* 1992). Most of these astronomy applications using ANNs have been reviewed by (Miller 1993).

In astronomy applications so far, researchers have used mainly two types of learning algorithms - a *supervised* backpropagation network (Rumelhart *et al.* 1986) and an *unsupervised* self-organizing map (Kohonen 1990). Hybrid models using both strategies, with and without supervision, have also been developed, such as the counter-propagation network (Hecht-Nielson 1987), which are not yet employed in astronomy. There are other models like hierarchical networks and tree networks (Sankar & Mammone 1991), which are still away from the astronomy field. Supervised learning requires researcher's desired response during training, which means, these networks learn in supervision of the researcher. Unsupervised learning requires no exemplar output or supervision.

For this application, we have investigated a conventional backpropagation network (hereafter, MBPN) as a pattern classifier to discriminate the IUE stellar spectra into various classes. We have also attempted a tree-like network, we call it as a multilevel tree network (hereafter, MTN), which employs the MBPN at the root and each non-terminal nodes of the tree.

The multilayer backpropagation network (Rumelhart et al. 1986) has several nodes arranged in a series of layers- an input layer, one or more hidden layers and an output layer. Input layer equals the size of the input pattern and the size of output layer is equal to the number of classes to be distinguished. The size of the intermediate layers depends on the complexity of the problem to be solved. The intermediate layers are also known as hidden layers. The network complexity is a function of the number of inputs, hidden nodes, layers, outputs and connections (see Fig. 1). The weighted connections between the layers are unidirectional and the information flows from the input layer through the output layer. Hence it is also called a feedforward network. A layer is fully connected if each node in the layer is connected to all the nodes in an adjacent layer. However the complexity of the network can be reduced by using partially connected layers. More detailed treatment on MBPN can be seen in the book on *Fundamentals of Neural Networks Architectures, Algorithms and Applications*, by L. Fausett.

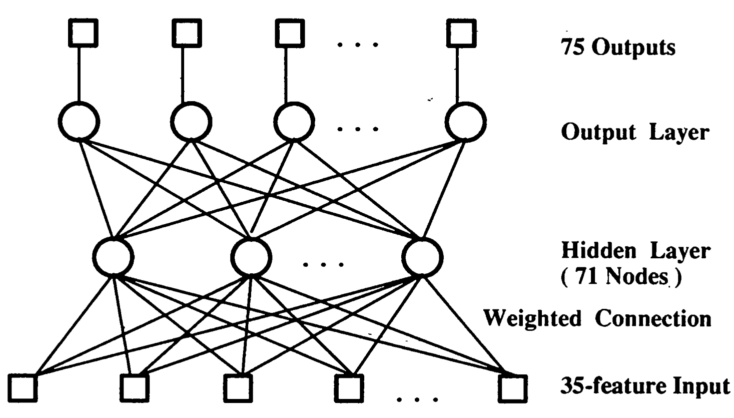

Figure 1. A fully connected backpropagation network with one hidden layer. Nodes applying activation functions are shown as circles. Non-contributing nodes are shown as squares which represent the 35-feature inputs and the 75 outputs.

In this paper, we have investigated two methods of ANN for the classification of IUE stellar spectra. The first method (ANN) involves a single MBPN classifier to classify the test spectra into 75 spectro-luminosity classes. The second method, which in principle incorporates a 'divide-and-conquer' technique, is devised as a multi-level tree network (MTN). The root node distinguishes the database into four main spectral types and then its children nodes further classified it into a total of 75 spectro-luminosity classes. Detailed description of this method is given in G94b.

4. Results and discussion

We applied ANN comprising the 35 input nodes (n=35 monitoring frequency parameters), one hidden layer of (2n+1) nodes and outputs of 75 nodes representing the 75 spectro-luminosity classes to train the network. The (2n+1) criterion is dictated by certain empirical facts and theoretical considerations which render an optimization on accuracy with (2n+1) neurons contained in the input layer (Murtagh 1991; Geva & Sitte 1992). A standard MBPN algorithm was used which consists of a training session followed by a testing session as described earlier. The inputs and outputs were connected by the usual sigmoid transfer function (G94a) with the initial inter-connection weights being generated randomly. The error in each iteration was compared with a set of desired output pattern and then back-propagated for updating the inter-connecting weights. Once the error in the network output was reduced to a pre-assigned threshold value, the iterations were stopped and the weights were frozen. The resultant frozen

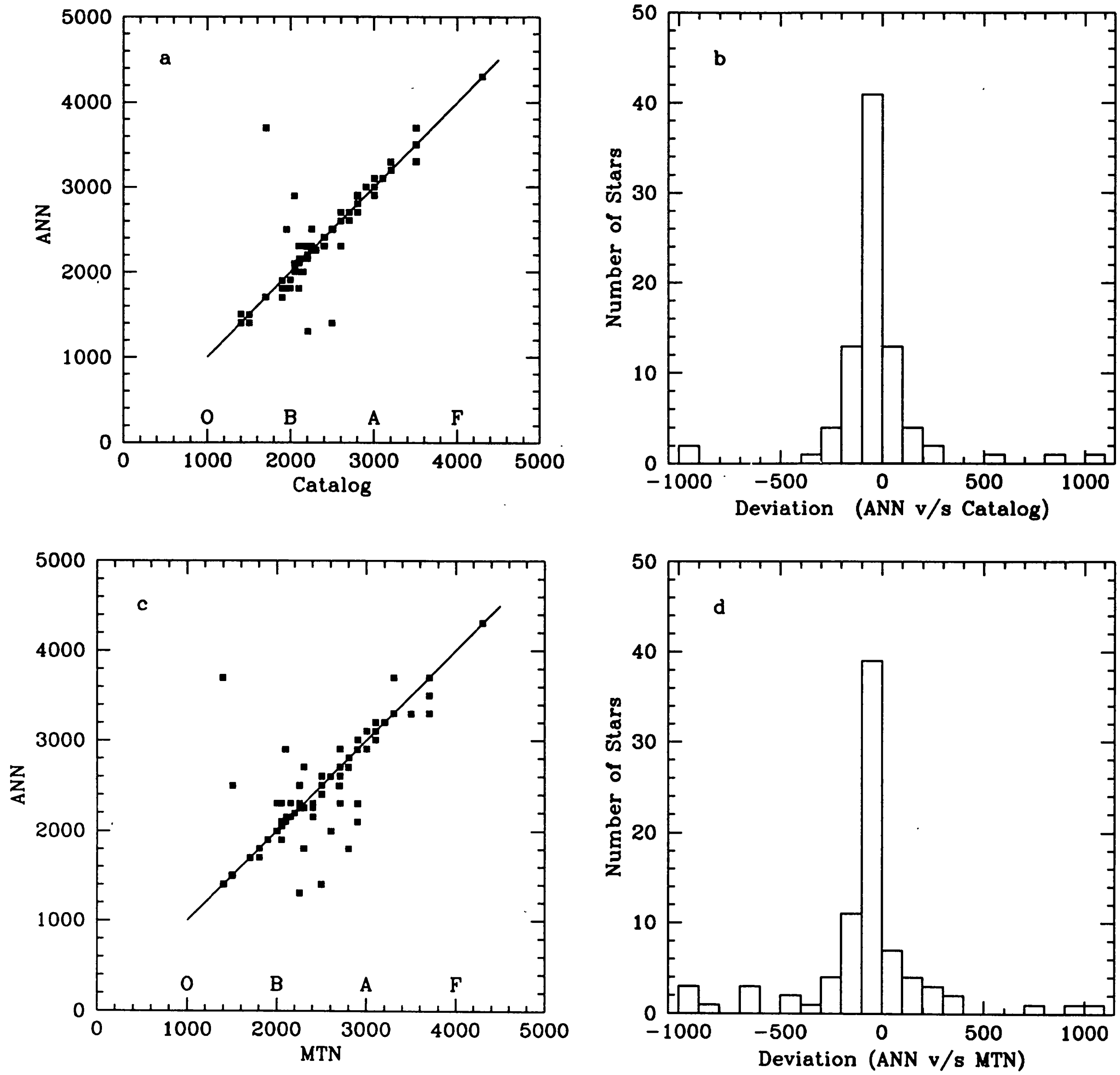

Figure 2. Scatter plots and corresponding histograms of deviations for: ANN vs. catalog in (a) and (b); ANN vs. MTN in (c) and (d). The diagonal line shows the ideal classification and the histogram bins correspond to 100 units of spectral code with respect to this line.

noting here that the advantage of using MTN over ANN is that the former is more efficient during training, once it has learnt the main classes in the root node level since it has to run only through a sub-set of the original data set. However it will be interesting to see the effect of various parameters involved in the model (Bhattacharya *et al.* 1995)

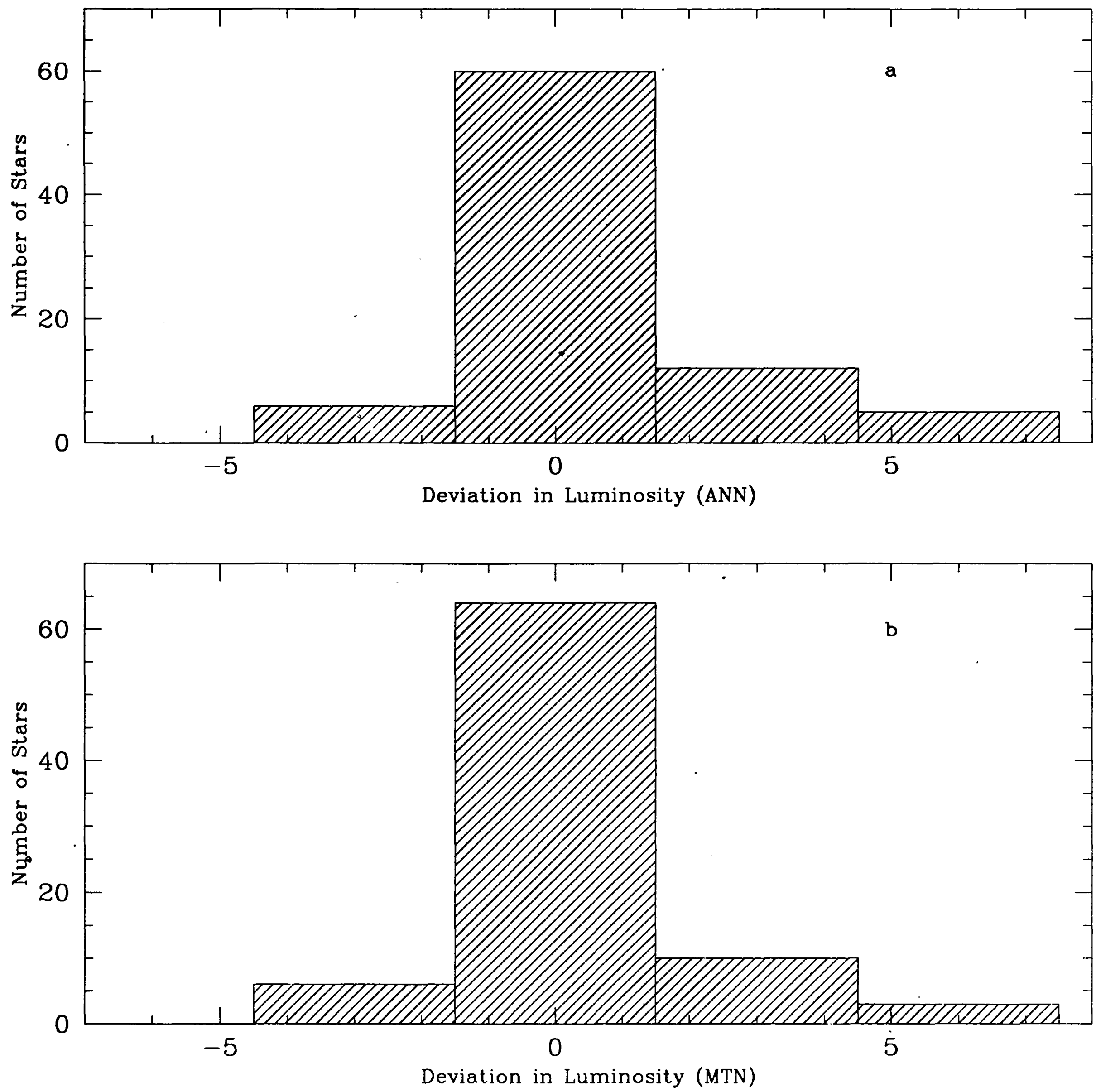

Figure 3. Histograms for luminosity classification in terms of deviations from the catalog luminosity for ANN in (a) and MTN in (b). Each bin corresponds to 3 luminosity units.

It may be noted that the values for A3 in eqn.(1) (the luminosity class) get comparatively lower weight when plotted with the four-digit-code number in the scatter plots of Fig. 2. To check the consistency of only the luminosity class based on MTN and ANN schemes, we separately investigated only the luminosity class and found that about (72%) of the total sample were classified correctly by these schemes. Figs. 3a-3b show the histogram of luminosity classification in terms of deviations from the catalog luminosity for ANN and MTN schemes respectively. As mentioned earlier we have restricted to three luminosity classes (super-giants, giants and dwarfs, i.e. 2,5 and 8), thus these histograms are binned at intervals of 3 luminosity units each. Thus our schemes can classify the IUE spectra in both dimensions viz. spectral and luminosity classes.

We further found that the UV and MK classifications differ by about 0.82 sub-classes and since this value is less than our best classification accuracy (~1.06 sub-classes), we cannot distinguish between UV and MK classifications. Possibly one may require higher spectral resolution for differentiating the two classifications. This is in agreement with similar conclusion arrived by Rountree & Sonneborn (1991). Finally, it is beyond the scope of this paper to investigate the individual outliers since here we are establishing the overall performance of the network schemes.

5. Conclusion

We have demonstrated that Neural Networks schemes based on multi layer back propagation method using single and multi-level tree configurations are able to classify most of the test sample of UV spectra to one-sub class accuracy. With these schemes 72% of the spectra are classified correctly in terms of luminosity which is otherwise a sore point of automated schemes. Further improvement is required in the training part of the ANNs which should ultimately lead to greater classification accuracies. ANNs will become extremely handy for large database classifications which are otherwise outside the scope of human classification. The requirements for ANNs would be to produce two-dimensional spectral classification with lowest possible spectral resolution, insensitiveness to interstellar reddening and with minimum human intervention. Unsupervised ANNs are expected to render such performance since our methods presented here still depend on the human classifier inputs.

We thank the IUCAA Astronomical Data Center for providing the spectral catalogs used in the current work. Thanks are also due to Profs. C. Jaschek and M. Jaschek for their useful comments. This research has made use of the SIMBAD database operated at CDS, Strasbourg, France. Finally, we wish to thank the anonymous referee for his constructive comments and suggestions.

References

Adorf H.M., 1986, in Data Analysis in Astronomy II, V, eds Di Gesu, et al., Plenum Press, NY, USA, p.61.

Adorf H.M. 1989 in Knowledge-based systems in astronomy, eds A. Heck & F.Murtagh, Springer-Verlag, Heidelberg, Germany, p. 215.

Adorf H.M., Johnston M.D., 1988 in Proc. Workshop Konnektionismus, eds C. Lischka 329, Kindermann J., Arbeitspariere der GMD, Bonn, Germany, p.3.

Adorf H.M., Meurs E.J.A., 1988 in Large Scale Structures in the Universe: Observational and Analytical Methods, eds W.C. Seither, H.M. Duerbeck & M. Tacke, Springer-Verlag, Heidelberg, Germany, p.315.

Bevington P.R., 1969 in Data Reduction and Error Analysis for the Physical Sciences, Chapter 1-5, McGraw-Hill, NY, USA.

Boggess A. et al., 1978a, Nature, 275, 372.

Boggess A. et al., 1978b, Nature, 275, 377.

Clark J.W., 1991, Phys. Med. Biol. 36(1), 1259.

Fausett L., 1993, Fundamentals of Neural Networks Architectures, Algorithms, and Applications, Prentice Hall, Englewood Cliffs, NJ 07632

Geva S., Sitte J., 1992, IEEE Transactions on Neural Networks 3, 621.

Gulati R.K., Malagnini M.L., Morossi C., 1987, JA&A, 8, 315.

Gulati, R.K., Gupta, R., Gothoskar, P., Khobragade, S., 1994, ApJ, 426(1), 340 (G94a).

Gulati R.K., Gupta R., Gothoskar P., Khobragade S., 1994, Vistas in Astronomy, 38(3), 293 (G94b)

Hecht-Nielsen (1987) in Proc. First Int. Conf. on Neural Networks, 2, p.19, IEEE 87TH0191-7, June, San Diego, CA, USA.

Heck A., Egret D., Jaschek M., Jaschek C., 1984, IUE Low-Resolution Spectra: A Reference Atlas, Part I, Normal Stars. ESA SP-1052.

Heck A., Egret D., Nobelis P.H., Turlot J.C., 1986, Astrophys. Space Sci. 129, 223.

von Hippel T., Storrie-Lombardi L.J., Storrie-Lombardi M.C., 1994 in The MK process at 50 years : A powerful tool for astrophysical insight, Sept.20-24, 1993, Tucson, Arizona, eds C.J. Corbally et al., A.S.P. Conf. Series, San Francisco, California, USA, 60, p.289.

Hopfield J.J., Tank D.W. 1986, Science, 223, 625.

Jaschek M., Jaschek C., 1990 in The MK Process and Stellar Classification, ed. R.F. Garrison, David Dunlap Observatory, Toronto, Canada. p.290.

Kohonen T. 1990, IEEE Proc. 78(9), 1464.

Kurtz M.J., 1984, in The MK Process and Stellar Classification, ed. R.F. Garrison, David Dunlap Observatory, Toronto, Canada, p.136.

Lahav O., Storrie-Lombardi, 1994, in Special Issue Neural Networks in Astronomy, Vistas in Astronomy, vol.38 (3).

McCulloch W.S., Pitts, W. 1943, Bull. Math. Biophys 5, 115.

Miller A.S., 1993, Vistas in Astronomy 36, 141.

Mukherjee S., Bhattacharyya U., Parui S., Gupta R., Gulati R.K., 1995, IUCAA Preprint 45 / 95.

Murtagh F. 1991, Neurocomputing 2, 183.

Odewahn S.C., Stockwell E.B., Pennington R.L., Humphreys R.M., Zumach W.A., 1992, AJ 103 (1), 318.

Rountree J., Sonneborn G., 1991, ApJ, 369, 515-528.

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

Sankar A., Mammone R.J., 1991, in Neural Tree Networks, eds R.J. Mammone, Y. Zeevi, Academic Press, NY, USA, p.281.

Storrie-Lombardi M.C., Lahav O., Sodre Jr.L., Storrie-Lombardi L.J., 1992, MNRAS, 259, 8.

Vieira E.F., Ponz J.D., 1995, AA Suppl. Ser, 111, 393.

Weaver Wm.B., 1994, in The MK Process at 50 years; A powerful tool for astrophysical insight, Tucson, Arizona, eds et al., C.J. Corbally, A.S.P. Conf. Series, 60, San Francisco, California, USA, p.303.

Weaver Wm. B., Torres-Dodgen A.V., 1995 ApJ, 446, 300.

Widrow, B., Lehr, M.A., 1990, in Proc. IEEE 78(9), p.1415.