Mem. S.A.It. Vol. 77, 1130

Memorie della

Automated classification and stellar parameterization

Sunetra Giridhar[1], S. Muneer[2], and Aruna Goswami[1]

[1] Indian Institute of Astrophysics, Bangalore 560034, India
[2] Indian Institute of Astrophysics, VBO, Kavalur, India

Abstract. Different approaches for automated spectral classification are critically reviewed. We describe in detail ANN based methods which are very efficient in quick handling of the large volumes of data generated by different surveys. We summarize the application of ANN in various surveys covering UV, visual and IR spectral regions and the accuracies obtained. We also present the preliminary results obtained with medium resolution spectra (R ~ 1000) for a modest sample of stars using the 2.3 m Vainu Bappu Telescope at Kavalur observatory, India. Our sample contains uniform distribution of stars in temperature range 4500 to 8000 K, log g range of 1.5 to 5.0 and [Fe/H] range of 0 to −3. We have explored the application of artificial neural network for parameterization of these stars. We have used a set of stars with well determined atmospheric parameters for training the networks for temperature, gravity and metallicity estimations. We could get an accuracy of 200 K in temperature, 0.4 in log g and 0.3 dex in [Fe/H] in our preliminary efforts.

Key words. Stars; automated classification; neural network

1. Introduction

The MK spectral types are classical description of stellar spectra. Although the two dimensional MK spectral type (SpT) and luminosity class (LC) are related to temperature and gravity of a star, the SpT is not assigned based upon these parameters but use visual appearance of stellar spectra. The MK classification involves comparing spectra to be classified with those of classification standards of defined class. The advantage of MK system has been that it is model independent and works well even with spectra of modest resolution. N. Houk and her collaborators have done a monumental work of determining SpT and LC for all stars in HD catalog (~ 12,000 stars up to Vmag ~ 11) with RMS error of 0.6 in SpT and 0.25 in LC. This data has been used as reference for automated classifications (see for example von Hippel et al. 1994, Bailer-Jones et al. 1998).

Many large telescopes are now equipped with multi-object spectrometers enabling coverage of a large number of objects per frame for stellar systems like globular clusters. Ongoing and future surveys, and space missions would collect a large number of spectra for stars belonging to different components of our Galaxy. Such large volume of data can be handled only with automatic procedures which would also have the advantage of being objective and providing homogeneous data set most suited for Galactic structure and evolutionary studies. Another outcome would be detection of stellar variability and finding of peculiar objects.

For a classification method to be effective, it should give precise SpT and LC for a wide range of spectral classes in a non-interactive way. The method should be precise enough not to assign earlier SpT to a metal-poor star.

2. Methods of automated spectral classification

The most commonly used automated spectral classification methods are based on (a) Minimum Distance Method (MDM) (b) Gaussian Probability Method (GPM), (c) Principal Component Analysis (PCA) and (d) Artificial Neural Network (ANN).

Quantitative methods involving measurement of equivalent widths of certain lines, line strength ratios, etc., and calibration of these quantities in terms of stellar parameters have also been used. For example, Stock and Stock (1999) used equivalent widths of 19 absorption lines, (B-V) colors and M_V derived from Hipparcos catalog for a sample of 487 stars for calibration of M_V. Their algorithms can predict M_V from these line strengths for spectral types O−M with an average error of 0.26 mag.

2.1. MDM

The classification is done by minimizing distance metric between the object to be classified and each member of a set of templates. The object is assigned the class of the template, which gives the smallest distance. If the star spectrum is represented by a vector $X = (x_1, x_2 ..... x_i, ... x_N)$ and template c is represented by another vector $S^c = (s^c_1, s^c_2 ..... s^c_i, ... s^c_N)$, the distance D_c is evaluated

$$D_c = \frac{1}{N}\left[\sum_{i=1}^{i=N} w_i^{\{c\}} \left|x_i - s_i^{\{c\}}\right|^p\right]^{1/p}$$

where $w_i^{\{c\}}$ is the weight assigned to spectral element i of the class c. Spectrum X is assigned class c for which D_c is minimum. Weights are assigned to spectral elements based on their relative importance in determining the spectral class. In this approach the number of templates used to define subclasses limit the accuracy of classification. Interpolation can be made to make intra-class assignment.

Katz et al. (1998) used this method with χ^2 weighing on high resolution Elodie spectra using a large number of reference stars of known T_{eff}, log g and [M/H] to derive atmospheric parameters of target stars. These authors achieved accuracy of 86 K in T_{eff}, 0.28 in log g and 0.35 in [M/H]. Vansevicius and Bridzius (1994) used MDM with χ^2 weighing to estimate SpT and M_V from Vilnius photometric indices. An accuracy of 0.7 was achieved for SpT and 0.8 mag for M_V over spectral type range O5 to M5.

2.2. GPM

It is a statistical approach of multivariate classification explained very clearly in Bailer-Jones (2001). It was employed by Christlieb et al. (2002) to estimate T_{eff}, log g and [M/H] for stars observed in objective prism Hamburg/ESO survey with an accuracy of 400 K in T_{eff}, 0.68 dex in log g.

2.3. PCA

It is a method of representing a set of N dimensional data by means of their projection onto a set of optimally defined axes. Since these axes (Principal components) form an orthogonal set, a linear transformation of the data is achieved. Not all components are important. Components that represent large variance are important while those that represent least variance can be ignored and data set can be replaced by significant components alone therefore resulting in reduction of the data size. These compressed data sets are used as input for neural networks. Bailer-Jones et al. (1998) had demonstrated that precise calibration could be done using these compressed spectra and that the optimal compression also results in noise removal. Singh et al. (2006) have used a variation of PCA technique to restore missing

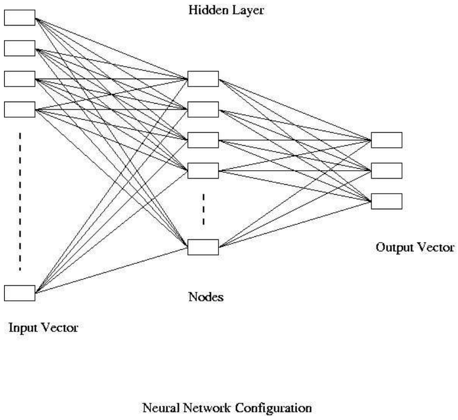

Fig. 1. A basic network configuration with three layers is presented.

data in a sample of 300 stars in Indo-US coudé feed spectral library.

2.4. Neural Network

As explained very well in numerous papers of Bailer-Jones, Ted von Hippel and others, it is a computational method which can provide non-linear parameterized mapping between an input vector (a spectrum for example) and one or more outputs like SpT, LC or T_{eff}, log g and [M/H]. The method is generally supervised, it means that for the network to give required input-output mapping, it must be trained with the help of representative data patterns. These are stellar spectra for which classification or stellar parameters are well determined. The training procedure is a numerical least square error minimization method. The training proceeds by optimizing the network parameters (weights) to give minimum classification error. Once the network is trained the weights are fixed, the network can be used to produce output SpT, LC or T_{eff}, log g and [M/H] for an unclassified spectrum.

As shown in figure 1, the neural network has one input layer containing stellar spectrum. Each of the input nodes connects to every node in the next layer of nodes called the hidden layer. The neural network architecture may contain one or more hidden layers. Each of these connections has a weight w associated with it. A given node in hidden layer forms a weighted sum of its inputs. It then passes this sum through a non-linear signoid transfer function to give final output from this node. The outputs from nodes in the hidden layer serve as input to the node in the output layer, which again forms weighted sum of its inputs. The training takes place as follows. The weights are initially set with random values over a small range. When the spectrum is fed into the network, the output would also be random. By comparing this output with the target output we can adjust the weights to give an output that is closer to the target value. The network is trained iteratively by successive passes of the training data through the network and on each pass the weights are perturbed towards their optimal value. The network training is performed by minimizing the least square error. Since the output from neural network is some non-linear function of all of the network input, it implies that the network output is based upon the appearance of whole spectrum. Depending upon the training data the network will learn which wavelength features are more significant than others in determining the correct spectral parameters and correspondingly would assign appropriate values to the network weights. Here the weights are updated backwords from the output layer through the hidden layer hence the algorithm is called back-propagation method.

3. Applications of the neural network in spectral classification

ANN has been used in a very large number of stellar applications. Vieira and Ponz (1995) have used ANN on low-resolution IUE spectra and have determined SpT with an accuracy of 1.1 subclass. although they attempted classification also with MDM, the errors of classification were larger than that of ANN. In visual region Bailer-Jones, Irwin and von Hippel (1998) used ANN to classify spectra from Michigan Spectral Survey with an accuracy of 1.09 SpT. Visual-near IR spectra were classified by Weaver & Torres-Dodgen (1997) using a two step approach. A first, coarse clas-

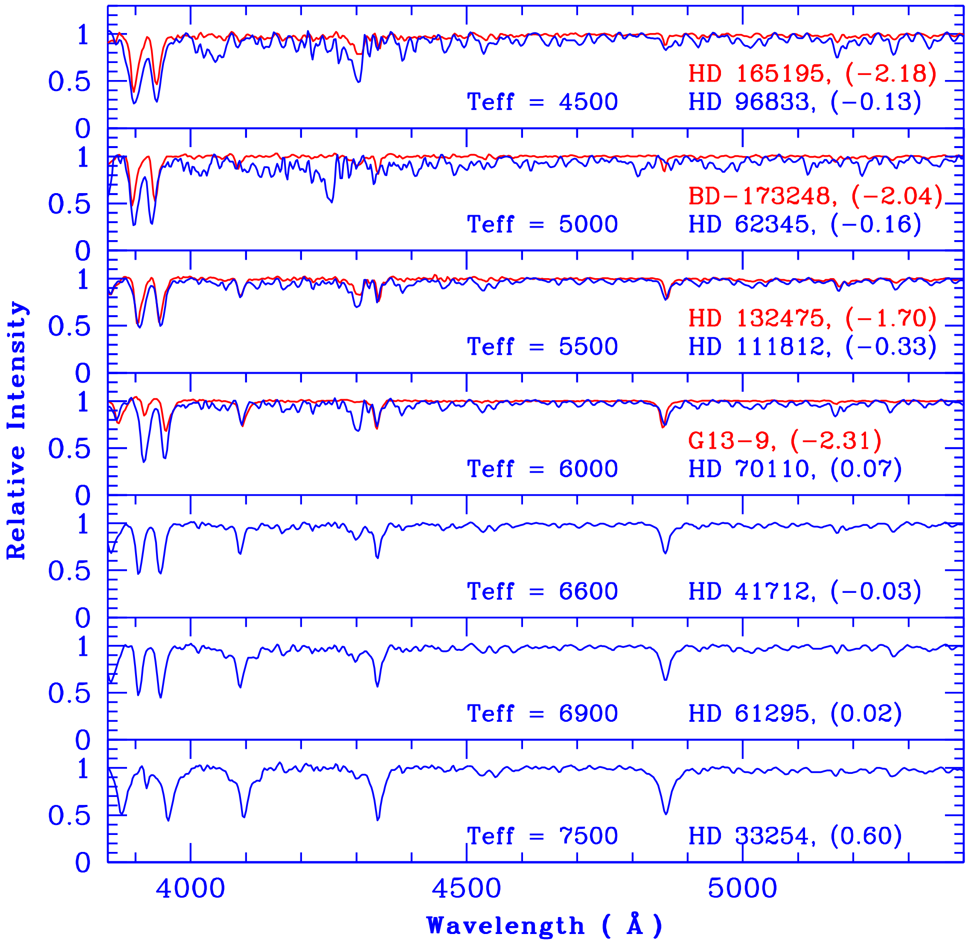

Fig. 2. A few spectra representative of the temperature range covered are presented. The metal-poor stars are plotted with red color.

sification is done to get the main spectral class, say F, then it is further classified by a more specialist network for that class. This approach results in an accuracy of 0.4 to 0.8 for SpT and 0.2 to 0.4 in LC.

Allende Prieto et al. (2000) used ANN in their search of metal-poor stars. Snider et al. (2001) used ANN for the three dimensional classification of metal-poor stars. Willemsen et al. (2005) have used ANN to estimate metallicity for main sequence turn-off, subgiants and red giant stars in the globular clusters M55 and ω Centauri using the medium resolution spectra of cluster members.

4. Parameterization of OMR Spectra

We have made a modest effort to use ANN for parameterization of a sample of stars in the temperature range 4500 to 8000 K. We have used a medium resolution Cassegrain spectrograph with the 2.3 m Vainu Bappu Telescope at

VBO, Kavalur, India. The spectrograph gives a resolution R (~ 1000) when used with a grating of 600 grooves/mm and a camera of focal length 150 mm. The spectra have a wavelength coverage of 3800-6000 Å. The pre-processing of spectra was carried out following a procedure very similar to that of Snider et al. (2001). We have observed stars from the list of Allende Prieto and Lambert (1999) and Snider et al. (2001) to develop a library of stars with known temperatures, gravities and metallicities. These spectra were used for training and testing the network. We have also observed stars from the lists of metal-poor candidates to estimate the metallicity for them. Although more than 200 stars are observed, here we report results for 90 stars for which atmospheric parameters T_{eff} and log g are well determined, [Fe/H] was known for 47 of them. We have used a software developed by B.D. Ripley based on back propagation technique. Figure 2 shows a few representative spectra. The preliminary results based on 680:11:3 architecture are presented in figure 3. The RMS error for T_{eff} = 200K, [Fe/H] = 0.3 dex, and that of log g = 0.4 dex. We propose to experiment with an architecture containing two hidden layers instead of one to further reduce the errors of these estimated parameters.

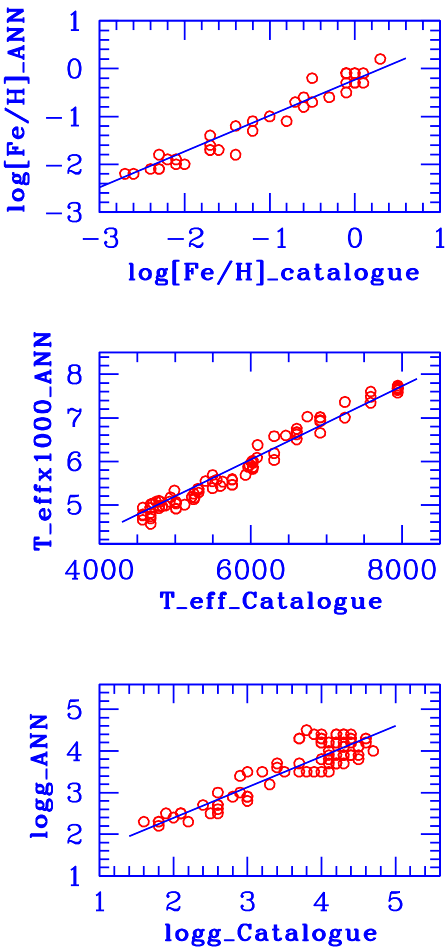

Fig. 3. The results from our preliminary analysis are presented. In the top panel we compare our ANN based estimates with those derived from high resolution analysis. A similar comparison is made for temperatures and gravities between our ANN based estimates and published results.

5. Future Goals

It is very important to envisage an approach that would give quick, reliable spectral classifications (or stellar parameters) for stars falling in all regions of the HR diagram. A single ANN architecture may not give the same desired accuracy over full range of spectral types and luminosity classes. A pilot program using photometric inputs, e.g. Strömgren indices or special photometric indices measuring the strengths of molecular bands for late type stars, could serve as preprocessor and help in identifying a set of specialist networks which would lead to classification of desired accuracy. A specialist system also needs to be evolved for A-type stars for a quick identification of chemically peculiar, magnetic or emission line stars. An expert system should also give strength of α elements or that of carbon using CH, CN bands.

Special network needs to be developed for objects displaying complex spectra such as symbiotic stars, novae and supernovae. Here the network must be trained on flux calibrated spectra and must use emission line strength as well as shape and structure of the continuum (composite for symbiotic stars and novae) for classification purposes.

Acknowledgements. It is a pleasure to thank Ted van Hippel for his help with software of B.D. Ripley and many tips towards the ANN architecture and its usage. Enthusiastic help from staff members of VBO Kavalur is also acknowledged.

References

Allende Prieto, C. & Lambert, D. L. 1999, A&A 352,555A

Allende Prieto, C., et al. 2000, AJ,120,1516A

Alonso, A., Arribas, S., Martnez-Roger, C. 1999,A&ASR, 140, 261A

Bailer-Jones, C.A.L. 2002,adaa.conf., 83B

Bailer-Jones, C.A.L., Gupta, Ranjan, Singh, Harinder P. 2002,adaa.conf, 51B

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

Christlieb, N., Wisotzki, L., Grahoff, G.2002,A&A, 391, 397C

Katz, D., Soubiran, C., Cayrel, R., Adda, M., Cautain, R.1998,A&A, 338, 151K

Singh, Harinder P., Bailer-Jones, C.A.L., Gupta, R. 2002,adaa.conf., 69S

Singh, Harinder P., Yuasa, M., Yamamoto, N., Gupta, R. 2006,PASJ,58, 177S

Snider, et al. 2001,ApJ, 562, 528S,

Stock, J. & Stock, J. M. 1999,RMxAA, 35, 143S

Vansevicius, V., Bridzius, 1994,BaltA, 3, 193V

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

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

Weaver, W., Bruce, Torres-Dodgen, Ana Vi. 1997,ApJ, 487, 847W

Willimsen,P.G., Hiker, M., Kayser,A., Bailer-Jones, C.A.L. 2005, A&A 436, 379W