**Automated Spectral Classification Algorithm Proposal**

**Rosie Fasullo**

**PHY 361 Advanced Research**

**August 26, 2018**

**Introduction**

Stars are grouped according to two different criteria, spectral type and luminosity class. The spectral type corresponds to the relative color of the light emitted by a star, which indicates the temperature of the star (Gray and Corbally 2009). There are 7 different spectral types used by astronomers today and they are, from to hottest to coolest, O, B, A, F, G, K, and M (Gray and Corbally 2009). Within each type there is a numeric subtype ranging from 0 to 9 with 0 corresponding to the hottest subtype and 9 corresponding to the coolest subtype (Gray and Corbally 2009). Luminosity class corresponds to the brightness of the star, which is an indication of the size of the star (Gray and Corbally 2009). Stars range from dwarf stars with luminosity class V to supergiant stars with luminosity class Ia (Gray and Corbally 2009). The current observing project at the Adams Observatory is the spectral classifications of O, B and A type stars, with the goal of better understanding their fundamental properties and types of chemical peculiarities.

There are several different systems used to classify stars. The system we employ is the MK classification system developed by W. W. Morgan and P.C. Keenan (Morgan and Kennan 1973). The MK system classifies stars by comparing spectrum of the star in question to a set of spectral standards and giving the star the label of whichever spectral standard it most resembles (Gray and Corbally 2009). The system does allow for the interpolation of spectral classifications if a star appears to lie between two spectra (Gray and Corbally 2009). However, this comparison is done using only spectra and does not admit the use of any extra information about the star such as photometric data (Gray and Corbally 2009). In this way, the MK system classifies stars according to their most natural groupings.

Although a few spectral classification algorithms have been created, the most reliable method is still classifying stars by eye. The MK classification system often relies on the astronomer’s intuition. Although computers are good at many things, they are unable to make decisions based on intuition and simulating intuition is very complex. Thus, it is usually easier and more accurate for spectral typing to be done by eye. However, given the number of stars in the sky and the time it takes to classify each star, an accurate and efficient spectral typing program would be hugely beneficial. What a team of astronomers could do over the course of several days, a reliable spectral classification program could do in a few minutes. Thus, we are creating an accurate and efficient spectral classification algorithm that will be able to classify normal as well as peculiar spectra.

**Approach**

We are going to use a data driven approach to creating a spectral classification algorithm. Each spectrum will have an associated list of characteristics that the program will keep track of, such as various metallic line strengths, metallic line shape, continuum shape and size, the presence of absence of emission, etc. The program will determine values for each of these characteristics when it is given a spectrum. To avoid consistency issues, we will begin with unrectified spectra and then preform an on the fly continuum rectification procedure for each one. Once each spectrum has been processed and labeled with the appropriate values for each of its characteristics, the data will be split into a training set and a test set. The classifier will be trained on the training set and then tested on the test set. All the spectra in both the training set and the test set will be pre-spectral typed by hand so that the accuracy of the algorithm can be easily determined. The classifier will train on the training set by identifying the trends or “natural groupings” of the data and creating an association between certain characteristics and the classification that matches those groupings of characteristics (Raschka 2015). Then, when given the test set, the classifier will group the testing spectra according to the associations it identified (Raschka 2015). We will test the accuracy of the classifier by having it compare the classification the program chose to the classification we assigned the spectrum.

**Objectives**

* Create a collection of spectral standards and stars that are very confidently spectral typed by hand that can be use
* Create a basic algorithm that can correctly spectral type the stars kept in the collection
* Expand the algorithm to be able to interpolate between spectral standards
* Implement functionality for the program to be able to identify stars that are not able to be accurately classified based on their spectra
* Expand the classification algorithm to cover spectra that are chemically peculiar
* Be able to detect and extrapolate individual spectra from a composite spectrum from a binary system

**Timeline**

|  |  |
| --- | --- |
| September 4, 2018 | Submit Research Proposal |
| October 30, 2018 | Submit Draft of Research Paper |
| November 12, 2018 | Oral Presentation |
| November 17-24 | Thanksgiving Break |
| November 29, 2018 | Submit Final Draft of Research Paper |

**References**

Morgan, W. W., and P. C. Keenan. “Spectral Classification.” *Annual Review of Astronomy and Astrophysics*, vol. 11, no. 1, 1973, pp. 29–50., doi:10.1146/annurev.aa.11.090173.000333.

Gray, R. O., and C. J. Corbally. “An Expert Computer Program For Classifying Stars On The Mk Spectral Classification System.” *The Astronomical Journal*, vol. 147, no. 4, 2014, p. 80., doi:10.1088/0004-6256/147/4/80.

Gray, Richard O., and Christopher J. Corbally. *Stellar Spectral Classification*. Princeton University Press, 2009.

Ness, M., et al. “The Cannon: A Data-Driven Approach To Stellar Label Determination.” *The Astrophysical Journal*, vol. 808, no. 1, 2015, p. 16., doi:10.1088/0004-637x/808/1/16.

Raschka, Sebastian. *Python Machine Learning*. Packt, 2015.