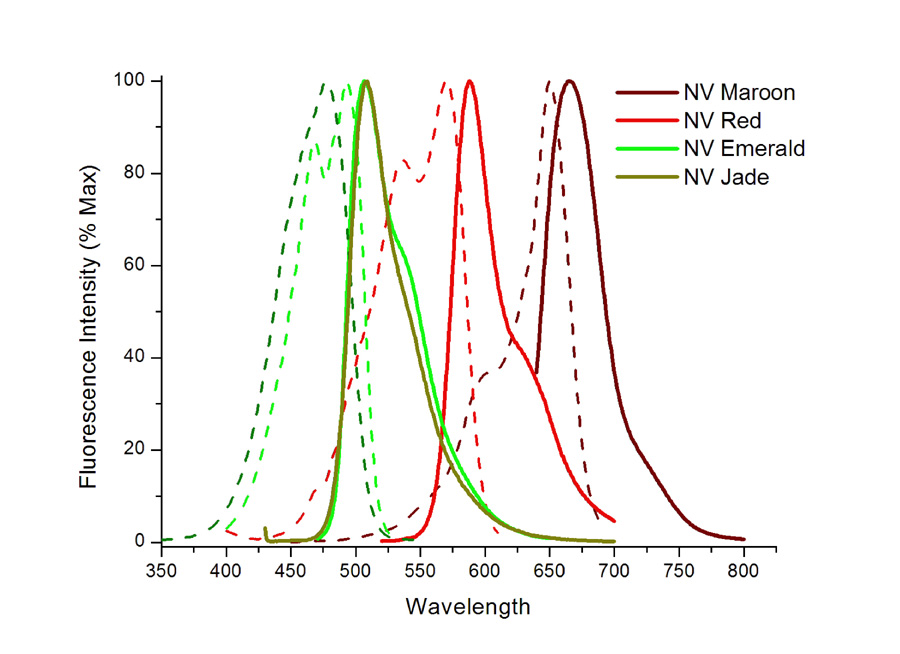
Prediction of Chemical Composition based on Fluorscence Spectra through the use of Neural Networks

For the course project, I would like to propose an application of a neural network (NN). The application involves a chemical/biochemical aspect to it as well, which I will briefly explain first. When performing a variety of biological studies, a common technique used is fluorescence. Fluorescence is a technique where a biological molecule (or any molecule) is labeled with another molecule (called a fluorophore). The fluorophore is a molecule that emits light at a specific wavelength, giving of a different color based on the chemical composition of the fluorophore. A fluorophore must first absorb energy, usually in the form of light, so that it can give off light to be measured by a detector. The detectors usually measure the intensity of the light given off by the fluorophore, usually in lumens. The detector then provides a spectrum of intensity vs wavelength for the fluorescence measurement. Currently, most spectra are simply a bunch of wavy lines, with various intensities, levels, etc. An example spectra is provided below:



Scientists have to go through and manually interpret the spectrum, see which peaks correspond to which fluorophores, or it is completely possible that the peak corresponds to the biological molecule itself because it is fluorescently active.

The project I propose is to create a neural network that will take a fluorescent spectra as input, and provide a prediction of what peaks correspond to which molecule. If a peak does not have a molecule, then an asterisk would be provided, marking that specific peak in the spectrum as one that should be further analyzed. Currently, fluorescent images of cells have been analyzed with neural networks, but normal spectra have not been [1]. The importance of this project would be to make the lives of scientists easier by providing an automatic means to analyze fluorescence spectra, removing the need to manually observe and identify peaks in the spectrum. This would in turn speed up the productivity in research, allowing more time for analysis and experimentation, rather than spending a bulk of time in analysis. Commonly, the coordinates will be provided for the graph, so the actual image analysis is not necessary.

I am currently looking for any neural network/software implementations that provide a skeleton that I can build upon to create the project, but have not been successful thus far. I plan writing my own software from scratch. The goal is to use some form of GUI that allows for dataset upload/spectrum upload, and then using a backend to perform the analysis. I am leaning towards using Erlang to create the neural network simply because Erlang was made for distributed processing, and the idea of lightweight neurons in a neural network is attractive. The dataset that I will use is any spectrum data I can get my hands on from scientists in the University of Maryland College Park, UMBC, as well as various data online. For the training data, I will use an online fluorescence database which contains various spectrums, giving me intensity, fluorophore identity, and wavelength of absorbance/emittance [2]. This should be enough for classification, but I will of course find other features to use as well as the project progresses. I currently work alone on the project, but if a team mate would like to join, I will let the Dr. Tim Oates know as soon as that happens. I hope to make advances in this project, and take it to a level outside of the classroom.

References

1. **Automated antinuclear immunofluorescence antibody screening: A comparative study of six computer-aided diagnostic systems**. Nicola Bizzaro, Antonio Antico, Stefan Platzgummer, Elio Tonutti, Danila Bassetti, Fiorenza Pesente, Renato Tozzoli, Marilina Tampoia, and Danilo Villalta. **Autoimmunity Reviews** (2014) 13: 292
2. **Spectra at UA**. http://www.spectra.arizona.edu/