Suggestions from Marvin da Silva

To use the Cov-TasNet paper (1D convolution) over the other 3 potential papers (LSTM): This was a great suggestion, I used the Cov-TasNet paper and I'm very glad I did. Their approach made a lot of sense, was very helpful, was able to be modified for my problem, and works well.

To use both raw spectral data and derived multiplet data as model inputs: I didn't implement this yet because I didn't have time, but I am going to.

Suggestions from Dr. Oore

To use a simplified synthetic dataset: I did this and it made a big difference. Initially I was going to use real data right away– this would have been a huge mistake because the database (single compound) and experimental (mixed compound) data weren't ready and the real data is large and complicated, which would have made figuring out the model very difficult. I gained a lot from making a synthetic dataset that is the same as the real data but very simplified. Having an initial model that works with the simplified synthetic data means I can more-or-less scale up the model when I use the real data.

To use source-separation: I did this and it was a great suggestion and works very well. Single-channel source separation, which is used for spectrograms (continuous data), also works very well for sequences (discrete data). The fact that source separation is meant for spectrograms means that using spectral data along with the derived multiplet data as inputs will likely work well.