7/05/2016

Today's Goal Finish progress report. Start experimenting with Supervised.

Unsupervised has so far been mostly unsuccessful due to the features not doing well with identifying REM. For now, I will be experimenting with supervised classification. For this, it will require some supervision of the data for 1-3 hours before automating.

I will be taking the features with F values >1000. This resulted in the following five features: Delta, Delta/Theta, EMG Power, Large/Small Ratio, and Sign Inversions.

I attempted to run a PCA with 2 comps. Below were the ANOVA results:

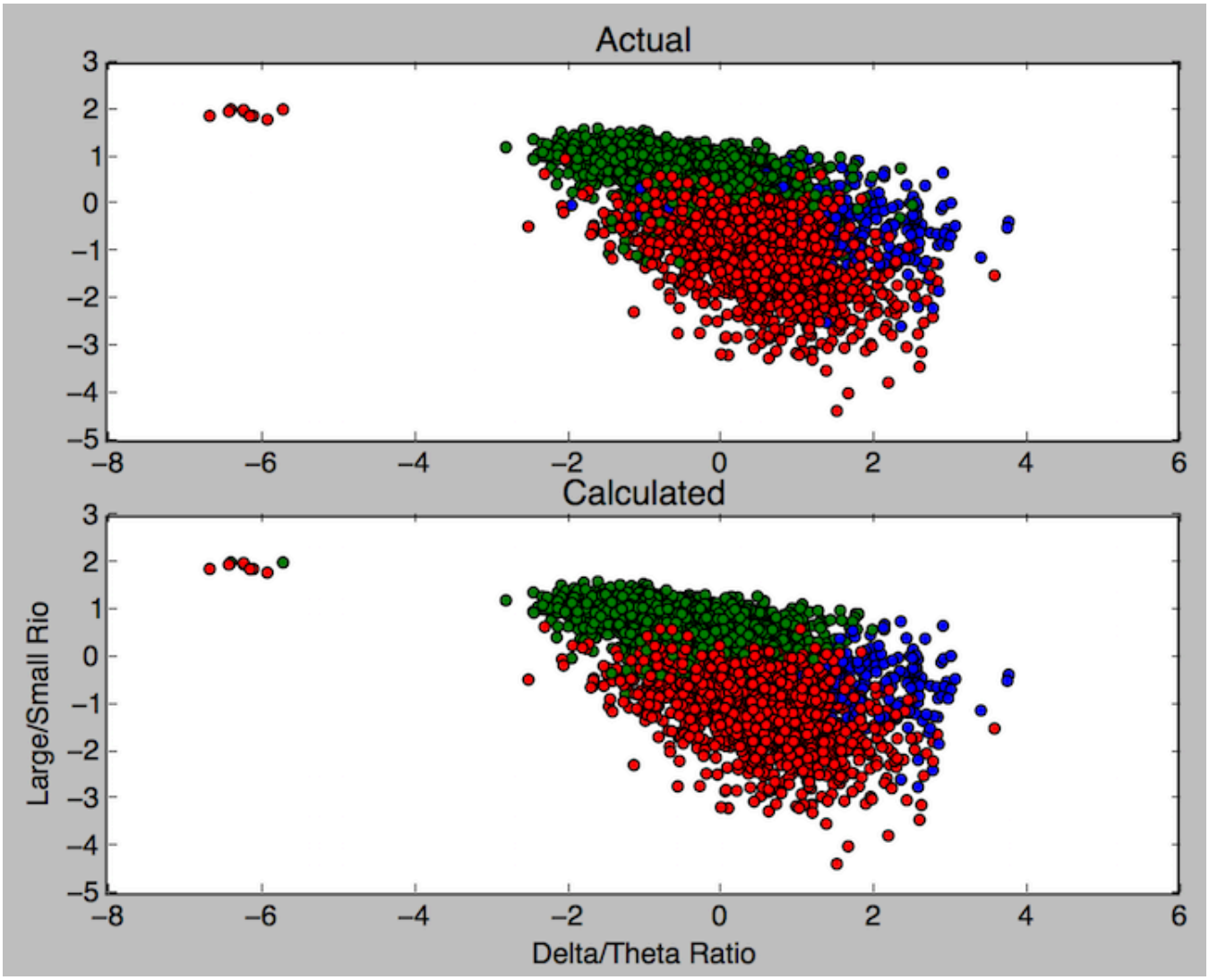
```
Feature delta has F=4630.7001365937485 and p=0.0
Feature delta/theta has F=551.7603980086735 and p=1.5545207680928752e-214
```

The PCA is probably not worth it, because the components have F-values that are worse than some of the original features.

I started experimenting with using SVM on subject 1032. Taking the whole data as training, here were the statistics:

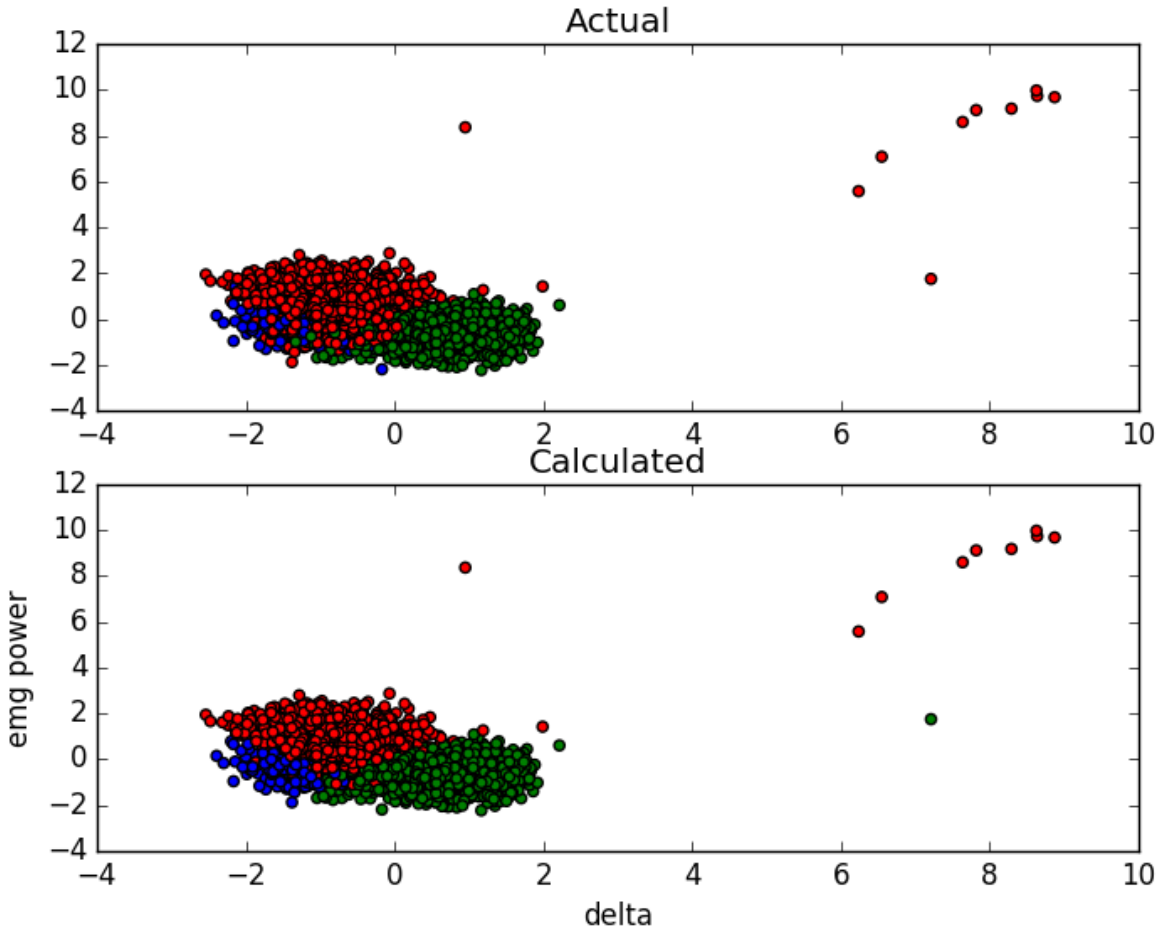
```
Wake. Calculated 1774 with actual 1792
NREM. Calculated 2443 with actual 2382
REM. Calculated 185 with actual 228
```

And here was one of the plots:

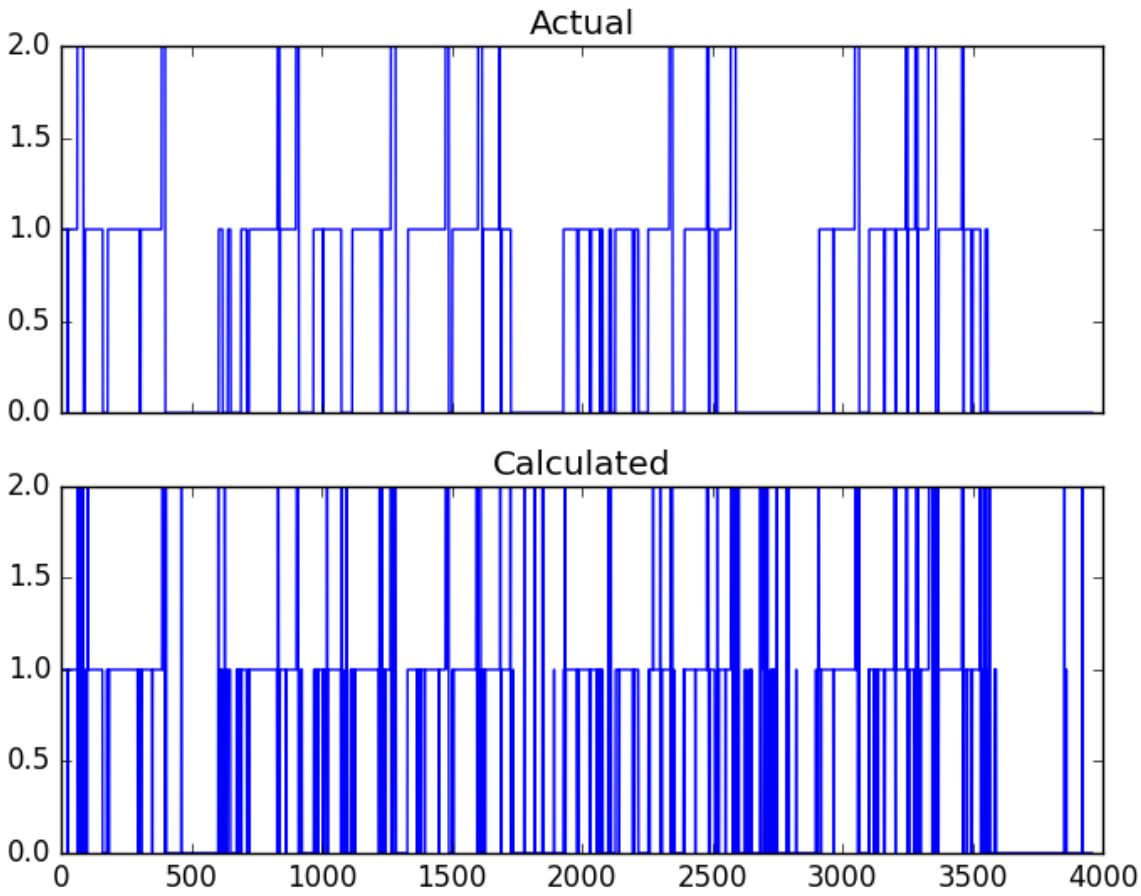


The Linear SVM worked quite well, and it was very fast too.

I used the first 440/4402 epochs to train the linear svm. The resulting accuracy for the rest of the data was 92.4%. Here is the plot of the testing data:



The SVM did extremely well while being computationally efficient. The downside is that it will require a couple of hours of supervision. Here is the hypnogram:



Although there are some false positives, the cyclical nature of the states is still preserved and it detects most of the actual states. Now I'm going to have to switch to supervised at the moment, but I will be experimenting with multiple algorithms (Neural Net, SVM, Decision Tree, Naive Bayes) for the real-time stage to see which ones win in both accuracy and computational efficiency/cost.

For now, the software will first record an initial 2-4 hours of data, have it manually scored by an experimenter, and then run overnight with the fitted classifier in place.