Beta regression sounds as an interesting variant to try. Trying to wrap my head around it, especially combined with this part of the "analysis_and_tips" python notebook, which (if I get this right) says that models trained on the two extreme classes versus the rest

both generate more or less the same result, while models trained on any of the middle classes versus the rest

generate results that are alike, but negatively correlated with the models trained on the extreme classes.

The first and last class are highly correlated

corrs=numpy.corrcoef(logistic.predict_proba(df[features]).T) plt.imshow(corrs, vmin=-1, vmax=1, cmap="RdYlGn") corrs

Maybe beta regression, rather than logistic regression changes that picture. Interesting thought. Alternatively, maybe there is some way to break up the tournament into 5 class-vs-other-classes, and put the results back together. The whole "extremes are correlated" reminds me of my badtimes

and goodtimes

models, which are also almost p/1-p :-). Have to think more.