

Hi, I did the exactly same thing and shared it here:[Clustering Eras - #7 by sneaky](#). Currently I'm working with 5 strong clusters of eras. I tried several methods to predict which era fits in which cluster, but IMO it is not possible with tournament data.

I tracked the performance of each cluster by learning a model on each and uploading the models predictions. When I compared the performances with real data like S&P500, VIX, ... I found some patterns. But, I doubt we have enough eras to train it.

I did a lot of experiments on this, but never had time to format it and share it in a clear form.

This picture shows the 5 (4) clusters that I found. Each dot represents similarity between two eras. Y and X axis are sorted by clusters.

- Cluster 0, which is the first square of eras in the interval from 0 to +/-120 is cluster of unique eras, that do not fit to any other cluster.
- Cluster 1, the second square, is the cluster of eras where the market acts predictably, also it is the cluster of eras that are easy to predict and most of the models have high correlations when they occur.
- Cluster 2 and 4, the third square, originally there were only 4 clusters, but I extended them to 5 and this image is from before.
- Cluster 3, the last square, this cluster is special, because it is negatively correlated to cluster 1, 2, and 4. The eras within the cluster are the eras that the majority of the tournament models have struggle with.

So I trained a model on each cluster separately and these are their results:

[

Screenshot from 2022-05-11 00-51-36

1319×190 32.6 KB

](<https://forum.numer.ai/uploads/default/original/2X/1/186cbc06460ad697612c99134e27c227b1c85881.png>)

It was evaluated on eras from validation data (V3 version)

It is clear that cluster 2 strongly correlates with cluster 4 (expected), and every model that wasn't trained on the cluster 3 has negative correlation with eras from the cluster 3.

This trend I found in other models I trained. Whenever I have model that is significantly better at predicting eras from cluster 3, the correlation of eras from the other clusters goes down.