**Summarized from output result charts**

**Found a best seed (random\_state=25) which has the least 0 values:**

Only Random Forest has 0’s (RF gets lots of 0’s for all other seeds, always)  
SVM gets 0’s without resampling for all the seeds, always.  
Other than that, all 3 thresholds for all the models have values.

ran more rounds with different random seeds, found another one: seed 5 is good too, but same as seed 25, still RF has lots of 0’s and SVM has 0’s for no-resampling. Generated result charts for seed 5. All other seeds’ results contain more 0’s. Didn’t make charts for other seeds, but their output files are in the folders.

**Threshold:**

0.5 and 0.7 are very close, and better than 0.97. ( 0.5 is the best )

**Resampling:**

By looking at different resampling methods’ results, I found oversampling the cheaters helped. So I added 1 resampling method, by changing the Smote portion of “Both 1:1” from 0.4 to 0.8, which enlarged the oversampling rate. It turned out that “Both 1:1 with smote\_rate=0.8” is the best.

Undersampler Only is still the worst.   
  
best to worst order:  
( Both 1:1 with smote\_rate=0.8 > Both 1:2 > Both 1:1 with smote\_rate=0.4 > Smote Only 1:2 > Smote Only 1:1 > No Resampling > Undersampler )

Skip Random Forest.

**To compare before/after resampling:**

The results are not very obvious, but it still shows most models got better scores after resampling.