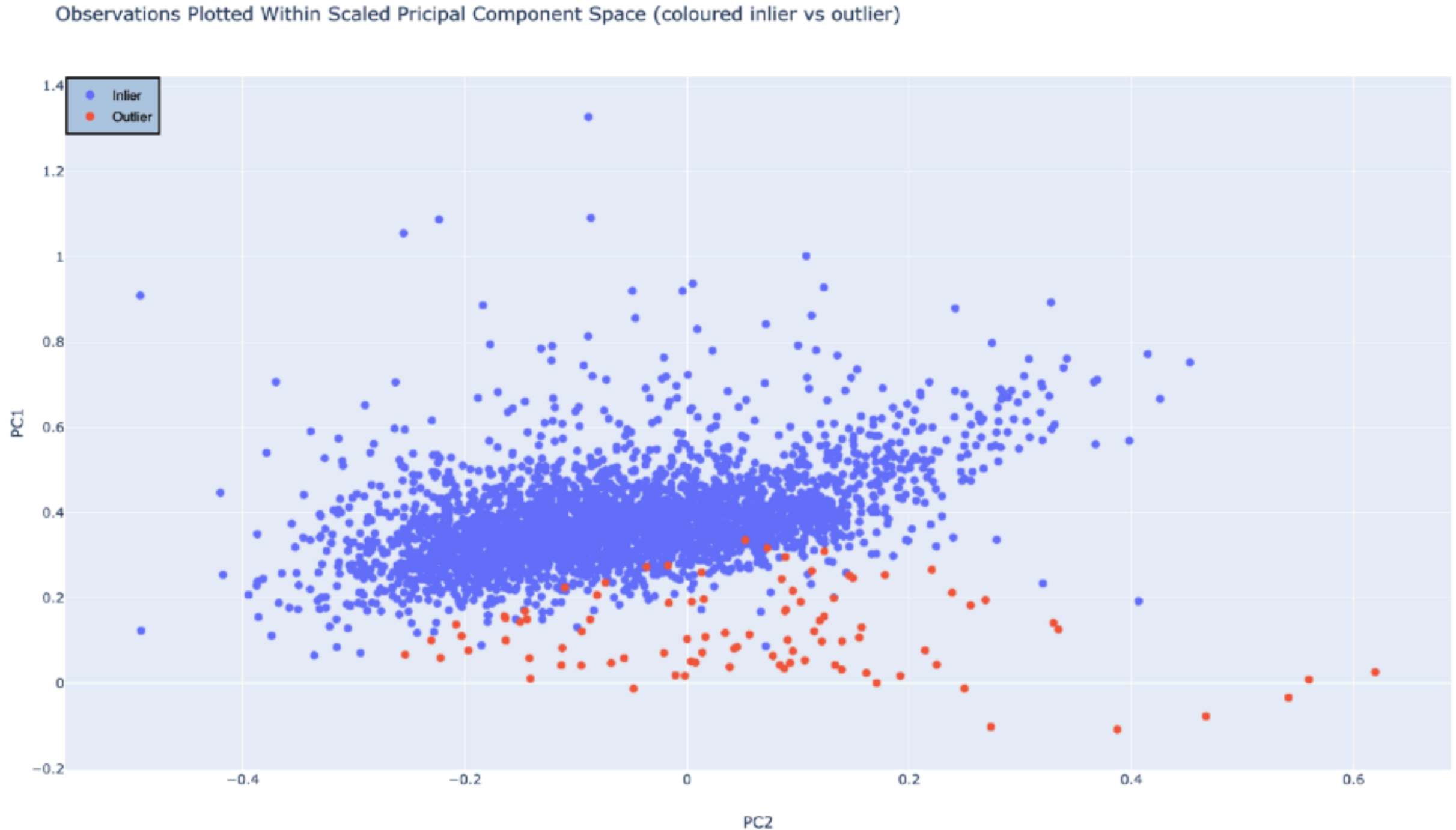


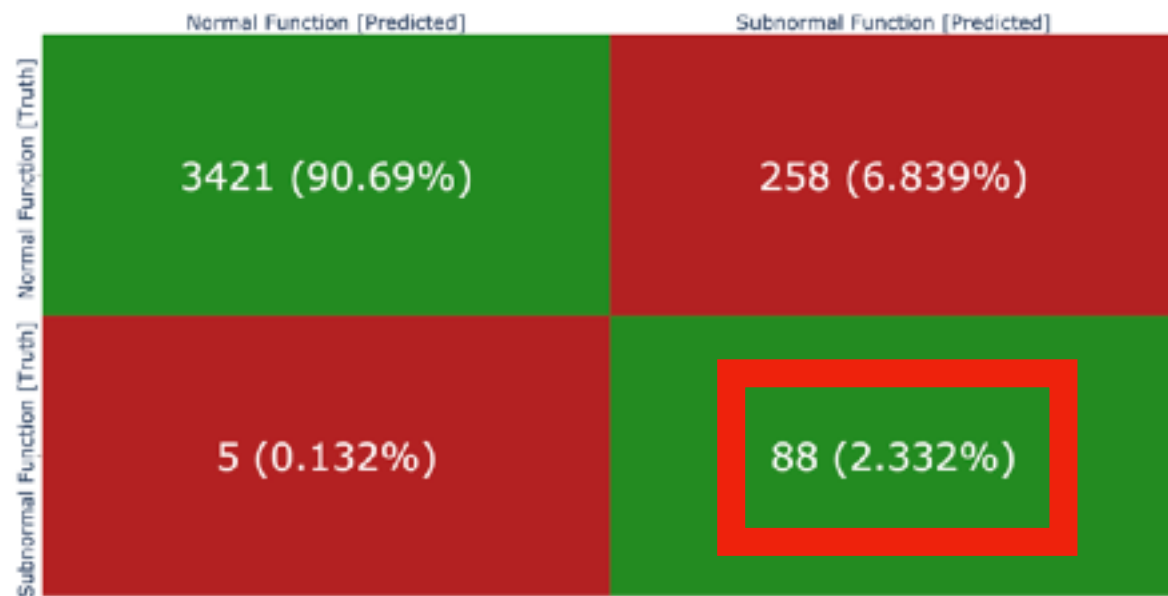
... but separation is good

(which is mostly what matters in this case)

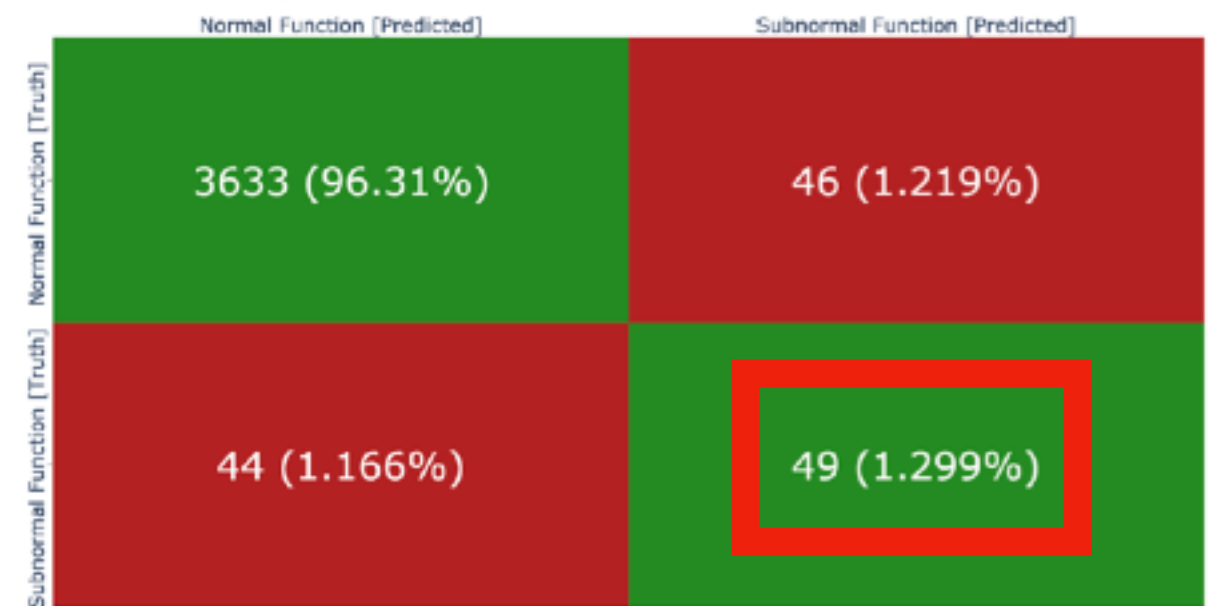


IF Performance

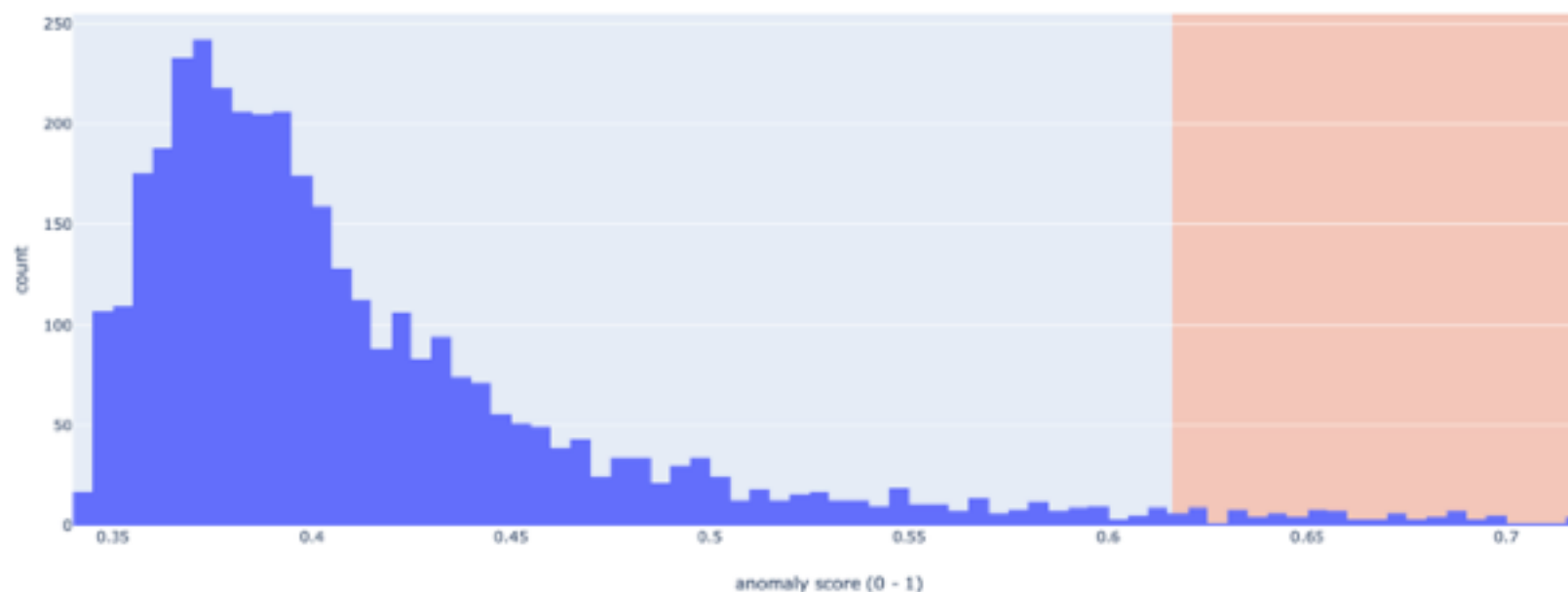
Isolation Forest (default) Confusion Matrix



Isolation Forest (contamination = 0.025) Confusion Matrix



Isolation Forest (contamination = 0.025) - Histogram of Anomaly Scores



Overall, not bad...

Threshold gets automatically set to a questionable point by default, but as soon as you set the correct contamination value the results seem respectable.

About half of the outliers within the set have been identified correctly as outliers. The other half being misclassified as inliers, and a set of similar size becoming false positives just to make life more difficult - could be worse.